Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop 10–11 March 1999

C.J. Pelley
Editor

9 July 1999

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LEXINGTON, MASSACHUSETTS

19990823 055

The ESC Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Gary Tutungian
Administrative Contracting Officer
Directorate of Contracted Support Management

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

PROCEEDINGS OF THE ADAPTIVE SENSOR ARRAY PROCESSING (ASAP) WORKSHOP 10-11 MARCH 1999

C.J. PELLEY
EDITOR

Division 10

PROJECT REPORT ASAP-7, VOLUME 1

9 JULY 1999

LEXINGTON   MASSACHUSETTS

1999 ASAP Workshop Theme

This year marks the seventh annual ASAP workshop, which is sponsored jointly by the DARPA Sensor Technology and Tactical Technology Offices, the Navy AEGIS and E2C Program Offices, and the NAVSEA Advanced Systems and Technology Office. This joint sponsorship reflects the far-reaching implications of adaptive sensor array processing for a wide range of military applications.

As this workshop has evolved over its seven year history, a common theme has been to highlight the similarity of adaptive sensor processing techniques for various disciplines. The ASAP forum has allowed researchers from academia, government, and industry to discuss common problems, developing ideas, and solutions related to how these diverse technologies can be applied to national defense interests. This year will continue to stress the strong coupling between the radar, sonar, and communications communities. A key theme will be the system implications of adaptive signal processing techniques. It is hoped that the cross-fertilization of ideas at ASAP will provide new areas of exploration and accelerate technological advances benefitting national defense interests.

MIT Lincoln Laboratory also hosts an important companion forum called the High Performance Embedded Computing (HPEC) Workshop which focuses on embedded computing architectures for real-time implementation of advanced adaptive signal processing techniques. This year HPEC will be held on 22–23 September 1999. The ideas presented in HPEC provide the essential link for allowing the advances highlighted in the ASAP workshop to be applied in real systems.

We welcome your participation in the seventh annual ASAP workshop.

Dr. Mark Davis
DARPA/ISO

Major Dan Gammon
DARPA/STO

Capt. Dan Meyer
US Navy/PMS-400
AEGIS

Capt. John Polcari
DARPA/TTO

Mr. Fred Lee
US Navy/PMA-231 AD

Mr. Jim Thompson
PEO Undersea Warfare Center

Workshop Chairman, Edward Baranoski
Technical Program Coordinator, Stephen Kogon
Workshop Administrator, Cindy Pelley

Technical Committee

Arthur Baggeroer / MIT
Lloyd Griffiths / George Mason Univ.
Peter Mikhalevsky / SAIC
Tom Miller / Raytheon
Kenneth Senne / MIT Lincoln Laboratory

Lee Swindlehurst / Brigham Young University
Donald Tufts / University of Rhode Island
James Ward / MIT Lincoln Laboratory
Michael Zatman / MIT Lincoln Laboratory

TABLE OF CONTENTS

PRESENTATIONS

PRESENTATIONS (Continued)

Circular Array STAP

Michael Zatman

MIT Lincoln Laboratory
244 Wood Street
Lexington, Massachusetts, 02420, USA

Abstract

Traditionally, Space-Time Adaptive Processing (STAP) has been applied to uniform linear arrays (ULAs). However, when considering the overall radar system, electronically scanned circular arrays have advantages: a better combination of even and continual angular and temporal coverage, and mechanical simplicity because it does not need to rotate. The unanswered question about circular arrays is their suitability for STAP.

This paper shows that for the airborne early warning (AEW) mission, circular arrays are indeed STAP compatible. However, from the STAP perspective there may be a small loss in performance when compared to a ULA. With some care in the choice and implementation of the STAP algorithm, the majority of the degradation is at close ranges, where the target returns are relatively strong. At long ranges the performance is barely affected.

1. Introduction

An experimental, circular, electronically scanned array is being fabricated by Raytheon as part of the Office of Naval Research sponsored UHF Electronically Scanned Array (UESA) program. The array has about 60 directional antennas arranged in a circular configuration. Here we assume that 20 of the 60 antennas will be used to transmit and receive at any one time. The antenna scans 'mechanically' in 6° increments by the choice of the 20 adjacent elements excited and 'electronically' +/-3°. Each antenna exhibits a front-to-back ratio of about 35 dB. The significant difference between the UESA and other airborne surveillance antennas is that UESA is a circular array, whereas the other antennas are uniform linear arrays.

In this paper the performance of conventional STAP algorithms on UESA-like circular arrays is assessed and compared to that of an 18 element ULA. The scenario assumes an AEW platform flying at a height of 9000 m, a velocity of about 100 knots and operating at 450 MHz with an 18 pulse CPI and a 300 Hz PRF. In Section 2 the theory of STAP is reviewed. The STAP performance of linear and circular arrays is assessed in Sections 3 and 4 for both full-dimension STAP and then reduced-dimension (pre-Doppler and post-Doppler) STAP respectively. In section 5 a technique for further improving circular array STAP performance is examined, and conclusions are drawn in Section 6.

2. STAP Fundamentals

STAP refers to the use of adaptive nulling techniques in the angle-Doppler domain to null clutter [1]. All forms of adaptive processing follow the flow of Figure 1, and perform best when the interference is statistically stationary. Thus STAP performs best when the locus of the clutter ridge is the same in the test cell and the training data.
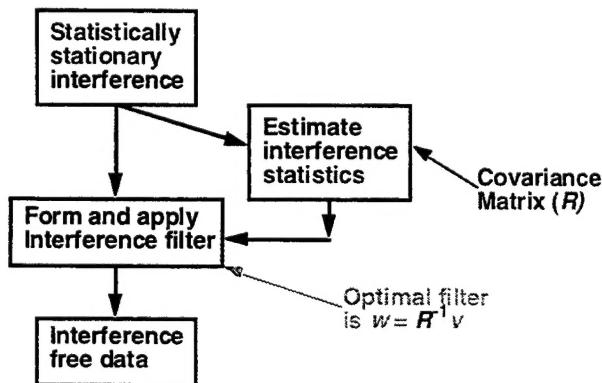
Figure 1: Adaptive signal processing chain.

Since the Doppler of clutter for a particular angle and elevation relative to the platform is given by

1

$$\frac{2 \times velocity \times \sin(azimuth) \times \cos(elevation)}{wavelength}$$

$$= \frac{2 \times velocity \times \sin(cone\ angle)}{wavelength} \qquad (1)$$

(i.e. the Doppler of the clutter is proportional to the sine of the cone angle of the clutter for a sideways looking linear array), the locus of the clutter ridge in the cone angle-Doppler space for a linear array is approximately stationary for most ranges of interest to airborne surveillance radars [2]. Thus, the assumption that the interference is statistically stationary is approximately true, and STAP works well for linear arrays.

Since a circular array sees two-dimensional angles, the clutter locus is no longer stationary. The azimuth and elevation of clutter with a particular Doppler change with range. The comparison between the angle-Doppler-domain view of clutter for both sideways looking linear and nonlinear arrays is shown in Figure 2. The change in the nonlinear array's clutter locus is fastest at close range.

Furthermore, because of the redundancy inherent in ULAs, the clutter rank is minimized for this configuration [3]. Since the element spacing of a circular array does not appear uniform when projected onto a linear aperture, the observed clutter rank will be higher than that of a ULA. Figure 3 shows the eigenvalues of the clutter covariance matrix at a range of 100 km for sideways looking (0°)

and rotated (30°) arrays. A clutter-to-noise ratio (CNR) of 45 dB per-element per-pulse is used. Note that for the sideways looking ULA the front and backlobe clutter fold on top of one another.
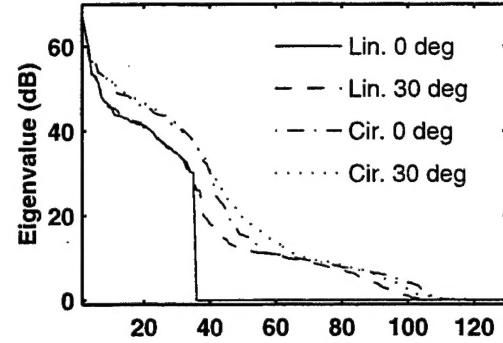
Figure 3: Clutter covariance matrix eigenvalues for linear and circular arrays with 0° and 30° of array rotation.

3. Full-Dimension STAP Performance

Although full-dimension STAP is not realizable for most practical AEW radars (because of the huge computational and sample support requirements), it gives a bound on the STAP performance that can be achieved by more practical algorithms. Figure 4 shows SINR loss [1] as a function of range (range³ CNR variation is assumed) and normalized Doppler frequency for the two arrays looking perpendicular to the aircraft's velocity vector (0° rotation).
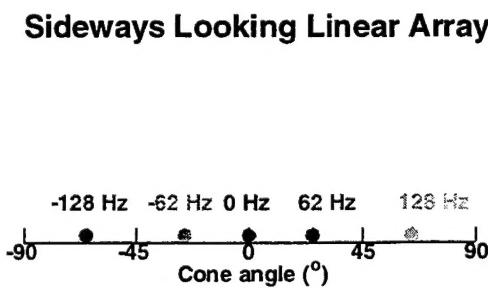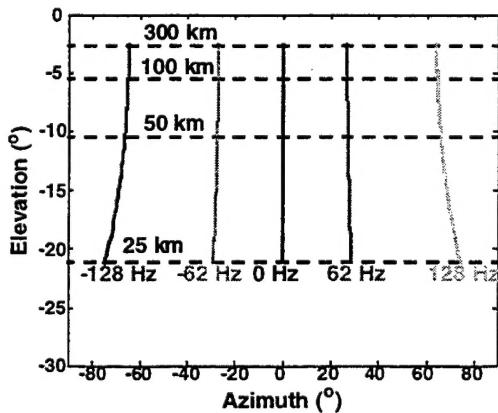
Figure 2: Angle of arrival (relative to platform broadside) for particular clutter Dopplers as a function of range. For the linear array the Doppler is a constant function of cone angle.

2
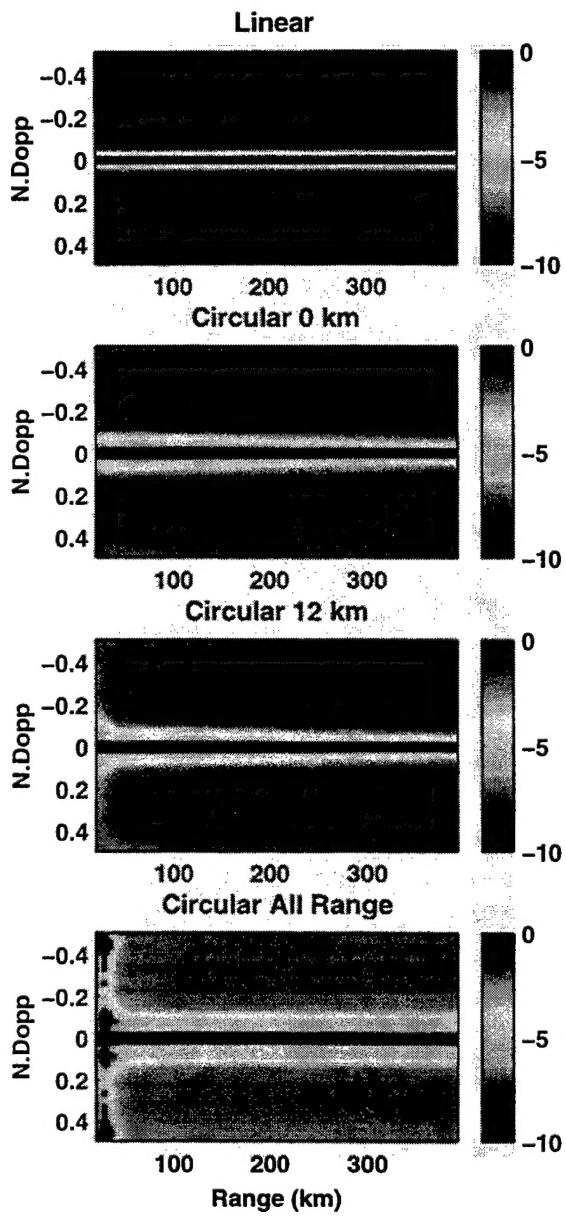
Figure 4: SINR loss (dB) as a function of range and normalized Doppler with no array rotation.

Figure 5: SINR loss (dB) with 30° of array rotation.

pronounced for the larger training windows due to the rapid change in the clutter locus with range. Close inspection also reveals that the best circular array SINR loss performance is slightly worse than that of the ULA (0.5 dB to 1 dB for 0 and 12 km training). The all range training shows a broad clutter null with particularly poor performance at close range.

For the ULA there is no change in SINR loss as a function of range (irrespective of the size of the training window). For the 0 km training region (impossible in practice) the width of the circular array's clutter notch broadens slightly at close range because of the higher CNR. This growth is more

Figure 5 shows the full-dimension STAP results with 30° of array rotation. The ULA's performance degrades because of the array rotation. The clutter null is wider (since more DOFs are now needed to null the clutter [1]). There is further degradation at close ranges when the effects of training included. With array rotation the locus of the clutter ridge for a linear array changes slightly as a function of range, though the effect is only noticeable at close range.

For the circular array the backlobe is not focused; thus, only one clutter null is visible. At close range the width of the clutter null grows in similar fashion to the cases depicted in Figure 4. With array rotation the overall performance of the ULA and circular arrays are similar.

4. Reduced-Dimension STAP Performance

Reduced-dimension STAP applies a dimension reducing transformation to the data prior to adaptive processing [1] and comes in pre-Doppler and post-Doppler, element-space and beam-space flavors. Here optimized pre-Doppler STAP [4] and PRI-staggered post-Doppler STAP [5], both in element space and with 3 temporal degrees of freedom, are compared.

Figure 6 compares the pre-Doppler STAP SINR loss performance of the linear and circular arrays with 30° of array rotation. The pre-Doppler STAP performance for a circular array with 6 km training is degraded from that of the linear array, slightly in terms of the SINR loss away from the clutter ridge (0.5 dB to 1 dB again), and the clutter null is significantly wider. The null width broadens slightly as the training window is enlarged to 12 km. For the all range training pre-Doppler STAP runs out of degrees of freedom, resulting in poor performance at all Doppler frequencies.

Figure 7 shows linear array pre- and post-Doppler STAP performance for a 6 km training window and post-Doppler circular array STAP performance for 6 and 12 km training windows. The well known result that post-Doppler STAP performs a little better than its pre-Doppler counterpart for the ULA is observed.

Unlike the pre-Doppler case, there is barely any growth in the width of the clutter null for post-Doppler STAP when switching from the ULA to the circular array. However, the circular array still suffers an additional SINR loss of about 0.5 dB away

from the clutter null. As expected, at close range the clutter null grows slightly as the training region is increased from 6 to 12 km. Comparison with Figure 6 shows that post-Doppler STAP performs significantly better than the pre-Doppler algorithm for the circular array.
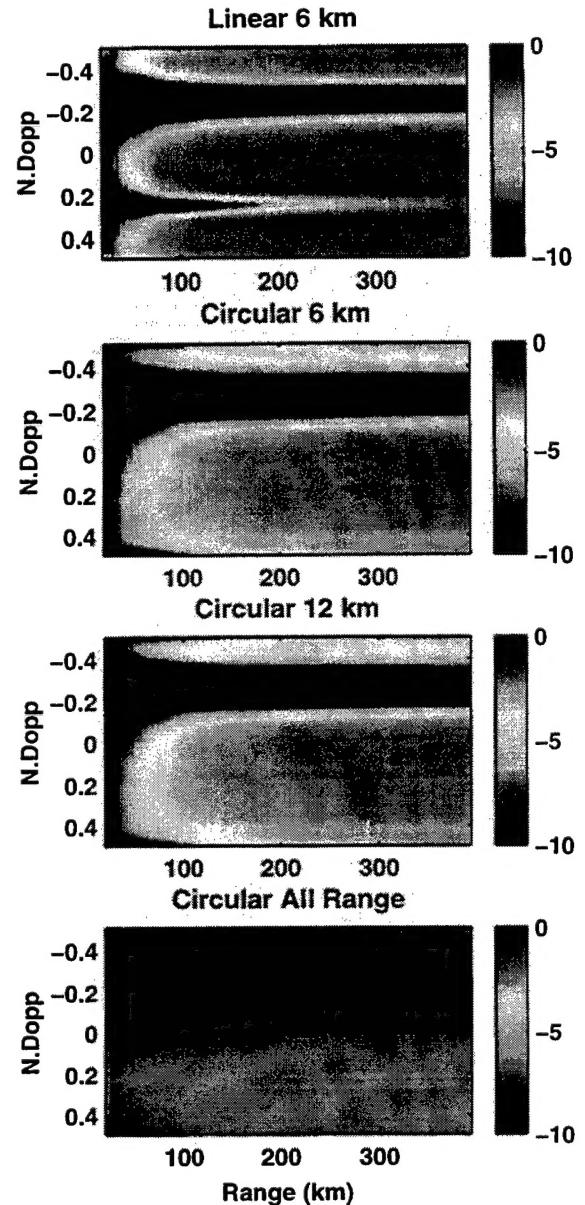
Figure 6: Comparison of pre-Doppler STAP performance on linear and circular arrays.

4
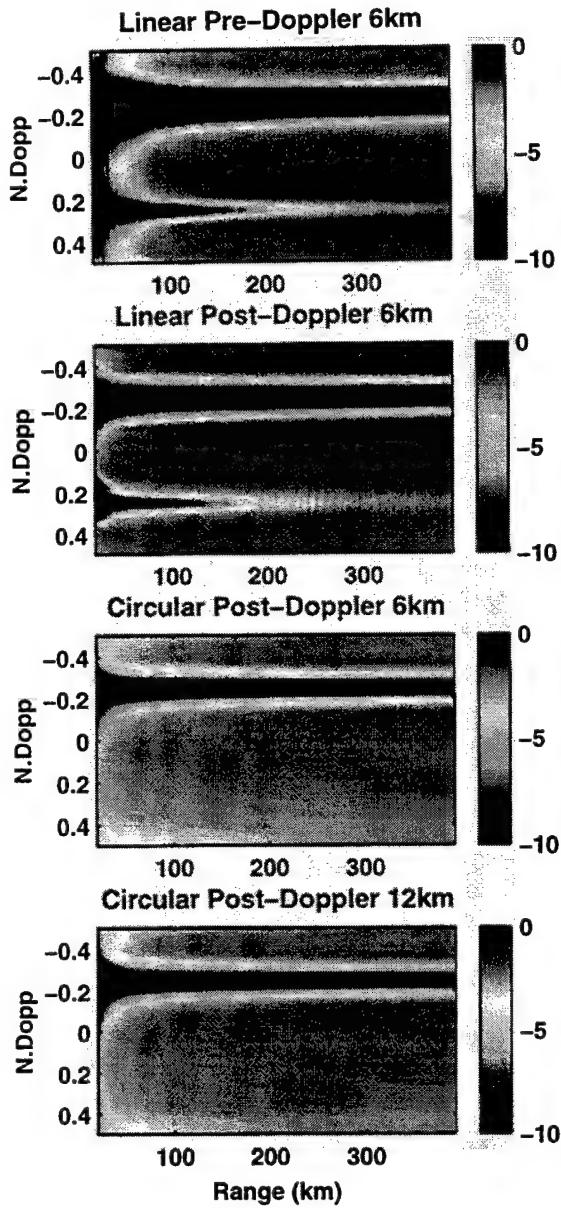
Figure 7: Pre- and post-Doppler STAP performance comparison with linear and circular arrays.

5. Range Varying Adaptive Weights

The non-stationarity of the clutter ridge as a function of range accounts for much of the degradation at close range. This problem is similar to that encountered when applying adaptive processing to a rotating array. The use of time varying adaptive

weights for rotating arrays with the extended sample matrix inversion (ESMI) technique [6] is widely applicable to a number of non-stationary problems including the case of the clutter ridge encountered in circular array STAP.

The instantaneous solution for the optimal adaptive weights as a function of range (r) may be expanded as the following power series

$$w(r) = w_0 + r\dot{w}_0 + \frac{r^2\ddot{w}_0}{2} + \dots \quad .$$

(2)

If the terms in r^2 and higher are small enough to be ignored, then for the kth range gate

$$w_k = w_0 + k\Delta w_0 \quad .$$

(3)

The two components of the time varying weight vector can be concatenated in the following form

$$w_x = \begin{bmatrix} w_0 \\ \Delta w \end{bmatrix} \quad ,$$

(4)

thus allowing the output of the range varying beamformer to be written as

$$y_k = w_x^H \begin{bmatrix} x_k \\ kx_k \end{bmatrix} = w_x^H X_k \quad ,$$

(5)

where x_k is the data vector at the kth rangegate, and X_k the corresponding 'extended data vector'. Thus an extended covariance matrix can be estimated from K snapshots of data

$$\hat{R}_x = \frac{1}{K} \sum_{k=1}^{K} X_k X_k^H \quad ,$$

(6)

and the adaptive range varying weights estimated as

$$w_x \propto \hat{R}_x^{-1} \begin{bmatrix} v \\ 0 \end{bmatrix} \quad .$$

(7)

Normalization of the weights (to say AMF or unit norm) and of X_k to ensure the proper convergence of the noise subspace eigenvalues is also required.

While the ESMI technique reviewed above takes account of the changing clutter locus, the size of the estimation problem is doubled, thus reducing the effective sample support. The post-Doppler STAP performance of ESMI at a range of 36 km is compared to regular SMI in Figure 8 with a 12 km training region for the adaptive weights. Also shown is the known covariance result assuming a 0 km training region. The ESMI technique produces a

narrower clutter null than regular SMI, but also incurs a small additional SINR loss away from the null due to the reduced sample support available. (The ESMI clutter null is slightly wider than the known covariance result due to finite sample size effects). Diagonal loading was used to improve the convergence of the adaptive weights.
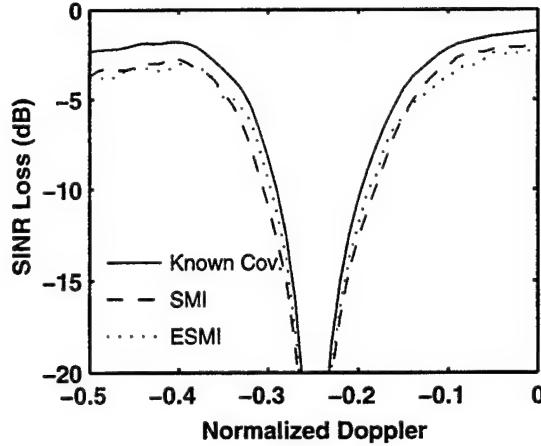
Figure 8: Post-Doppler STAP performance of the SMI and ESMI techniques.

One advantage of post-Doppler STAP is that the adaptive processing for each Doppler bin need not be identical. For circular array STAP this conveniently allows the region at close range adjacent to the main beam clutter to be processed with the ESMI technique, ensuring narrow clutter nulls and good minimum detectable velocity performance. The regular SMI technique is applied everywhere else, thus minimizing both the SINR loss away from the clutter and the computational burden.

6. Conclusions

This paper has shown that practical STAP with a UESA-like antenna is possible. There are two effects that lead to degraded performance relative to a uniform linear array.
- The clutter locus changes as a function of range, leading to broader clutter nulls and reduced radar minimum detectable velocity performance, because some range extent is required to train the adaptive weights.
- The clutter rank at a single rangegate is not minimized and the backlobe clutter is unfocused. This leads to a slight widening to the clutter null

and a small additional SINR loss away from the clutter null.

The first of the effects can be mitigated to some degree by the use of range varying weights.

The full-dimension STAP results show that the use of highly localized training regions for the adaptive weights results in acceptable circular array STAP performance. The null width only grows significantly at close range (as it does for a linear array with rotation), and the reduction in SINR performance at other Dopplers is only about 0.5 dB to 1 dB. Apart from the sweet-spot in uniform linear array performance when the array is aligned with the platform's velocity vector, circular array STAP performance is comparable with that achieved by the uniform linear array.

The post-Doppler STAP has similar performance characteristics to the full-dimension STAP, making it attractive for practical implementation. Furthermore, post-Doppler STAP performance can be maximized at close range by implementing range varying adaptive weights in the Doppler filters adjacent to the mainbeam clutter. Unfortunately, whereas pre-Doppler STAP performs well on uniform linear arrays, it degrades significantly when applied to a circular array.

References
[1] Ward, J., "Space-time adaptive processing for airborne radar," MIT Lincoln Laboratory technical report 1015, Dec. 1994.
[2] Brennan, L.E., Mallett, J.D. and Reed, I.S., "Adaptive arrays in airborne MTI radar," IEEE Trans. AP, Vol. 24, 1976, pp. 607-615.
[3] Ward, J., "Space-time adaptive processing with thinned arrays," Asilomar Conference on Signals, Systems and Computers, 1998.
[4] Baranoski, E.J., "Constraint optimization for pre-Doppler STAP algorithms," Asilomar Conference on Signals, Systems and Computers, 1998.
[5] Brennan, L.E. and Staudaher, F.M., "Subclutter visibility demonstration," Adaptive Sensors Inc. technical report RL-TR-92-21, 1992.
[6] Hayward, S.D., "Adaptive beamforming for rapidly moving arrays", Int. Conf. on Radar, Beijing, 1996.

Principal Components, Covariance Matrix Tapers and the Interference Modulation Problem

J. R. Guerci

Science Applications International Corp.
4001 N. Fairfax Drive, Suite 400
Arlington, VA 22203 USA
JRGuerci@ieee.org

J. S. Bergin

Information Systems Laboratories, Inc.
8130 Boone Boulevard, Suite 500
Vienna, VA 22182
JSB@isl-inc.com

ABSTRACT

In this paper, a new approach to adaptive interference mitigation is presented which combines the principal components (PC) method with the concept of a covariance matrix taper (CMT). The PC-CMT method is shown to preserve the minimal sample support property of the PC method even when there is significant eigenspectrum spreading of the dominant interference subspace. In addition to demonstrating its utility for STAP applications, a novel inverse CMT (ICMT) method is suggested for eigen-based direction finding applications for which signal modulation is an issue.

1. INTRODUCTION

Principal components (PC) methods [1] have several advantages when applied to many space-time adaptive processing (STAP) applications. When the interference is dominated by a strong low rank subspace, it has been demonstrated that a significant reduction in sample support requirements can be achieved [2]. For example, in the case of a uniform linear array (ULA) with $\lambda/2$ interelement spacing, in the absence of antenna "crab," Brennan [3] has shown that the clutter rank $K \approx \lceil N + \beta(M-1) \rceil$, where N is the number of antenna elements, M is the number of pulses processed in a coherent processing interval (CPI), $\lceil \ \rceil$ denotes the "ceiling" operator (round to nearest largest integer), and β is the distance traveled in a PRI measured in interelement spacing units [4]. For the DARPA Mountain Top radar [5], $N=14$, $M=16$, and $\beta \approx 1$, resulting in $K \approx 29$. Thus, for moderate to strong clutter environments [2], a sample support of only approximately $2K = 58$ space-time snapshots (range bins) is required to achieve an SINR that is within 3 dB of optimum (on average). This is in contrast to a full degree-of-freedom (DOF) based sample matrix inversion (SMI) method [6] that would require approximately $2NM \sim 448$ samples to achieve the same level of SINR performance. A case for which the clutter-noise-ratio (CNR) is 40 dB is shown in Figure 1. Note that essentially equivalent performance is achieved by the PC approach with significantly less sample

support. Other advantages of the PC method include excellent adapted pattern sidelobe control, compatibility with other rank-reduction and minimal sample support methods such as pre/post-Doppler [4] and F/B smoothing [7]-[8], less target signal cancellation (fewer "guard cells" required), as well as it's applicability to nonlinear (e.g., circular) arrays [9].

Unfortunately, actual real-world eigenspectrum are rarely observed to have a sharp delineation of the dominant subspace. There are in fact many mechanisms that can conspire to significantly increase the observed interference rank. Examples include (but are not limited to) nonlinear array geometries [10], ULA antenna "crabbing" [4], clutter scintillation/motion [4], bandwidth effects [11], diffuse multipath or antenna "jitter" [12]-[13], and transmitter instabilities (Doppler modulation). As a consequence, PC based methods must either increase the effective rank of the interference subspace, (with a commensurate or greater increase is sample support requirements), or, if they maintain the original rank, suffer an additional SINR loss due to interference undernulling.

In this paper, we exploit the fact that a large class of interference modulation effects is accurately modeled as a "tapering" of the unmodulated covariance matrix [11]-[13]. A novel PC-CMT approach to interference mitigation is introduced which is shown to restore the minimal sample support property of the PC approach, yet not suffer from undernulled interference. This is accomplished by first estimating the dominant eigenvectors (which are essentially unaffected by the modulation) utilizing a commensurately small sample support, then synthesizing the remaining subdominant interference subspace by applying a CMT.

The remainder of the paper is organized as follows: In Section 2, the connection between interference modulation and CMTs is established. The new PC-CMT interference mitigation algorithm is introduced in Section 3. Simulation results establishing the efficacy of the approach are presented in Section 4. In Section 5, implementation issues and other novel applications of the CMT approach are discussed. Section 6 contains a summary.
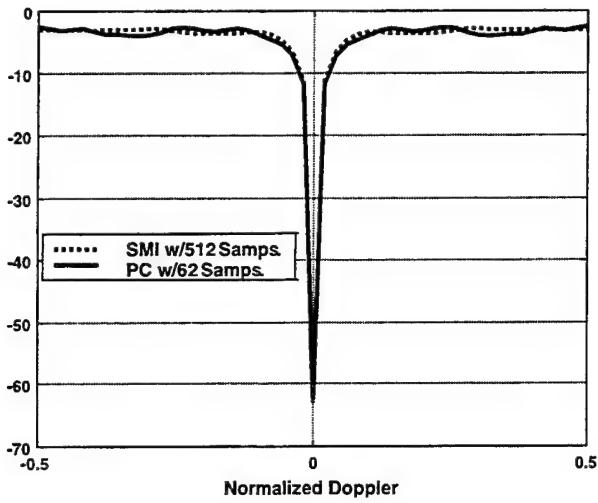
Figure 1 SINR loss (in dB) versus normalized Doppler for the PC and SMI methods.

Figure 2 CMT for ICM example with a motion parameter of 1 m/s.

2. INTERFERENCE MODULATION AND COVARIANCE MATRIX TAPERS

Let $\mathbf{x}_C \in C^{NM}$ denote the "structured" or "colored noise" portion of the total zero mean random vector process interference signal. Examples of relevance to STAP applications include clutter and jamming. Many mechanisms that result in an increase in the effective rank of the colored noise portion of the interference subspace are manifested as a random "modulation" of the form

$$\mathbf{x}_{CT} = \mathbf{x}_C \circ \mathbf{x}_T \in C^{NM} \qquad (1)$$

where $\mathbf{x}_T \in C^{NM}$ is a zero mean vector random process which is uncorrelated with $\mathbf{x}_C$, and $\circ$ denotes the element-wise multiplication operator (Hadamard Product [14]). It is relatively straightforward to show that

$$R_{CT} = \text{cov}(\mathbf{x}_C \circ \mathbf{x}_T) = R_C \circ T \in C^{NM \times NM} \qquad (2)$$

where $R_{CT} \overset{\Delta}{=} \text{cov}(\mathbf{x}_{CT})$, $R_C \overset{\Delta}{=} \text{cov}(\mathbf{x}_C)$, and $T \overset{\Delta}{=} \text{cov}(\mathbf{x}_T)$ is a CMT [13]. Thus, random modulations of the form (1), result in a covariance matrix tapering of the unmodulated interference covariance matrix [13].

An example of a CMT for internal clutter motion (ICM) which has a Gaussian correlation function [4] is depicted in Figure 2. We have assumed a N=10 element ULA with half-wavelength interelement spacing and a CPI of M=10. In Figure 3, the corresponding eigenspectra for several different "motion" parameters are shown. Note that it appears that the dominant eigenvalues are not affected by the modulation. Indeed, it is possible to show that the dominant eigenvalues/vectors are not appreciably affected by moderate modulation. Consider the Karhunen-Loeve (KL) expansion of the vector random clutter process

$$\mathbf{x}_C = \sum_{i=1}^{K} \tilde{\alpha}_i \mathbf{u}_i \qquad (3)$$

where $\mathbf{u}_i \in C^{NM}$ is the i-th eigenvector of R_C and $\{\tilde{\alpha}_i\}$ denotes the set of complex scalar KL expansion coefficients. Note that we have assumed that "Brennan's Rule" holds in (3), i.e., the KL expansion need only include the K dominant terms. A CMT modulation has the form

$$\mathbf{x}_C \circ \mathbf{x}_T = \left(\sum_{i=1}^{K} \tilde{\alpha}_i \mathbf{u}_i \right) \circ \mathbf{x}_T = \sum_{i=1}^{K} \tilde{\alpha}_i \mathbf{u}_i \circ \mathbf{x}_T . \qquad (4)$$

From (4) we see that the introduction of CMT modulation is tantamount, in general, to a random amplitude and phase modulation of the dominant eigenvectors. Thus for moderate modulation, the process is analogous to a communication system in which a strong carrier (eigenvector) is weakly modulated. The result is a power spectrum in which the carrier still dominates, but for which "sidebands" have been introduced. This relative invariance of the dominant eigenvectors serves as a basis for the new PC-CMT algorithm described in the following section.

8
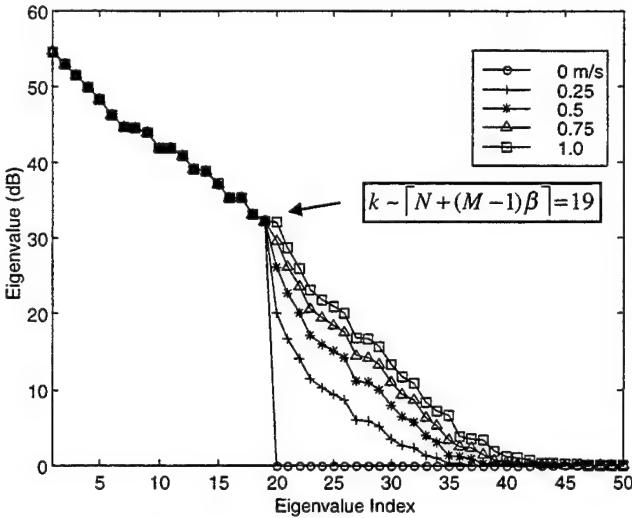
Figure 3 Eigenspectra for varying amounts of ICM. Note the relative invariance of the dominant eigenvalues.

Figure 4 Performance of the proposed PC-CMT approach as compared to traditional SMI with a 2NM sample support, PC by itself, and PC combined with the cross-spectral method (CSM).

3. COMBINED PC-CMT STAP

The invariance of the dominant eigenvectors suggests the following estimation procedure:

1. Estimate the K dominant eigenvectors using a PC method and sample support O(2K), i.e., employ PC as if no modulation were present.

2. Synthesize the subdominant eigenvectors by applying a CMT to the K dominant eigenvectors obtained in Step 1.

3. Obtain a full rank estimate of the total interference covariance matrix $R_I \in C^{NM \times NM} > 0$ by restoring the receiver noise floor.

The above procedure can be succinctly described by

$$\hat{R}_I = \left(\sum_{i=1}^{K} \hat{\lambda}_i \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i' \right) \circ T + \sigma^2 I \qquad (5)$$

where $\left\{ \hat{\lambda}_i, \hat{\mathbf{u}}_i \right\} i = 1, \ldots, K$ are the K dominant eigenvectors obtained from Step 1, T is a CMT, and σ^2 is the receiver noise. Based on this estimate, an optimal STAP beamformer is obtained in the usual manner

$$\mathbf{w}_{\text{PC-CMT}} = \hat{R}_I^{-1} \mathbf{s} \qquad (6)$$

where $\mathbf{s}$ is the desired target steering vector (angle-doppler) of interest [4].

4. SIMULATION RESULTS

Figure 4 compares the relative performance of the proposed PC-CMT approach with PC only, SMI with a 2NM sample support, and a PC method which employs the cross-spectral metric (CSM) [15]. The PC-CSM employs explicit knowledge of the desired target steering vector to form a more "judicious" rank ordering of the eigenspectrum, i.e., eigenvectors are ranked on the product $\hat{\lambda}_i (\hat{\mathbf{u}}_i' \mathbf{s})$ rather than simply $\hat{\lambda}_i$. Clearly evident is the near restoration of the O(2K) minimized sample support property of the PC-CMT approach.

Figures 5 (a)-(b) demonstrate the robustness of the PC-CMT approach to CMT mismatches. In fact, for very small sample supports O(K), there are advantages to applying a "heavier" taper to aid in synthesizing the underdetermined interference subspace.

5. IMPLEMENTATION ISSUES AND OTHER APPLICATIONS

5.1 Alternate Implementations

If an eigencanceler based STAP beamformer is employed, some measure of computational reduction can be achieved. Recognizing that the first K dominant eigenvectors of the full rank estimate of R_I given by (5) are well approximated by Step 1 of the proposed PC-CMT algorithm (see Section 3), one need only consider the remaining eigenvectors of $\hat{R}_I$ given by (5) in the null space of $\left\{ \hat{\lambda}_i, \hat{\mathbf{u}}_i \right\}$ obtained from Step 1. More specifically:

9
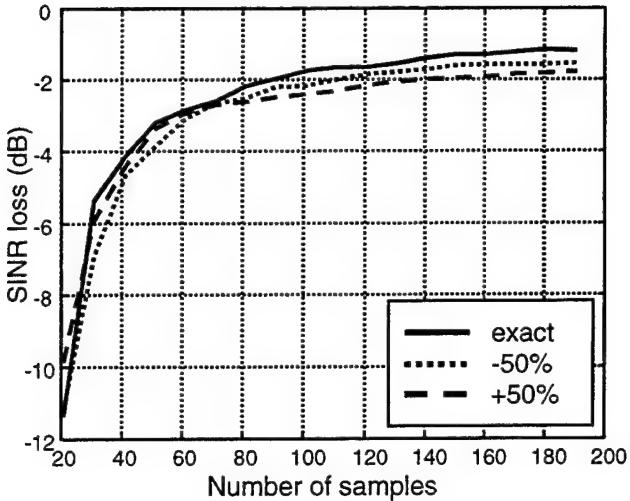
Figure 5 (a) Robustness of PC-CMT to +/- 50% uncertainty in CMT parameter (nominal is 1 m/s). Normalized Doppler = 0.125.

Figure 5 (b) Normalized Doppler = 0.25.

$$\mathbf{w}_{\text{PC-CMT}} \approx \sum_{i=1}^{K} \frac{\hat{\lambda}_i - \sigma^2}{\hat{\lambda}_i}(\hat{\mathbf{u}}_i' \mathbf{s})\hat{\mathbf{u}}_i + \sum_{i=K+1}^{L} \frac{\hat{\lambda}_i - \sigma^2}{\hat{\lambda}_i}(\hat{\mathbf{u}}_i' \mathbf{s})\hat{\mathbf{u}}_i \quad (7)$$

where the first K terms are obtained from Step 1 and the remaining terms are obtained by continuing the eigendecomposition from the $K+1$ stage. Another alternative is to approximate the dominant eigenvectors by 2D DFT beams pointed at the clutter ridge.

5.2 The Inverse CMT (ICMT) Method

Eigenbased direction finding algorithms [16] can be adversely affected when eigenspectrum spreading is present. Interestingly, the CMT modulation effect on the unmodulated covariance matrix can, in theory, be undone. For example, if only the modulated covariance estimate is available, i.e., $R \circ T$, then it is possible to eliminate T by performing a Hadamard inverse [14], i.e., $R = (R \circ T) \circ (1./T)$, where $1./T$ denotes the Hadamard inverse (element-wise inversion). This inverse CMT (ICMT) approach is illustrated in Figure 6.

6. SUMMARY

A novel PC-CMT method was introduced which was demonstrated to restore the minimal sample support property of the PC method even when significant interference modulation (and consequently eigenspectrum spreading) was present. Some thoughts on more computationally attractive implementations of the PC-CMT approach, along with other potential application areas, were also presented.

7. REFERENCES

[1] I. P. Kirsteins and D. W. Tufts, "Adaptive detection using low rank approximation to a data matrix," *IEEE Trans. on Aerosp. Electron. Syst.*, vol. 30, pp. 55-67, January 1985.

[2] C. H. Gierull, "Statistical analysis of the eigenvector projection method for adaptive spatial filtering of interference," *IEE Proc. On Radar, Sonar and Navig.*, vol. 144, pp. 57-63, 1997.

[3] L. E. Brennan and F. M. Staudaher, "Subclutter Visibility Demonstration," Technical Report RL-TR-92-21, Adaptive Sensors Inc., March 1992.

[4] J. Ward, "Space-Time Adaptive Processing," Technical Report 1015, MIT Lincoln Laboratory, December 1994.

[5] G. W. Titi, "An overview of the ARPA/Navy MountainTop Program," *Proceedings of the IEEE Adaptive Antenna Systems Symposium*, November 7-8, Long Island, New York, 1994.

[6] I. S. Reed, J. D. Mallett, L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. on Aerosp. Electron. Syst.*, vol. 10, no. 6, pp. 853-863, 1974.

[7] M. A. Zatman and D. Marshall, "Forwards-Backwards averaging for adaptive beamforming and STAP," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*," Atlanta, GA, 1996.

[8] S. U. Pillai, J. R. Guerci and Y. L. Kim, "Generalized forward/backward subaperture smoothing techniques for sample starved STAP," *Proceedings of 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Seattle, WA, May 12-15, 1998.

10
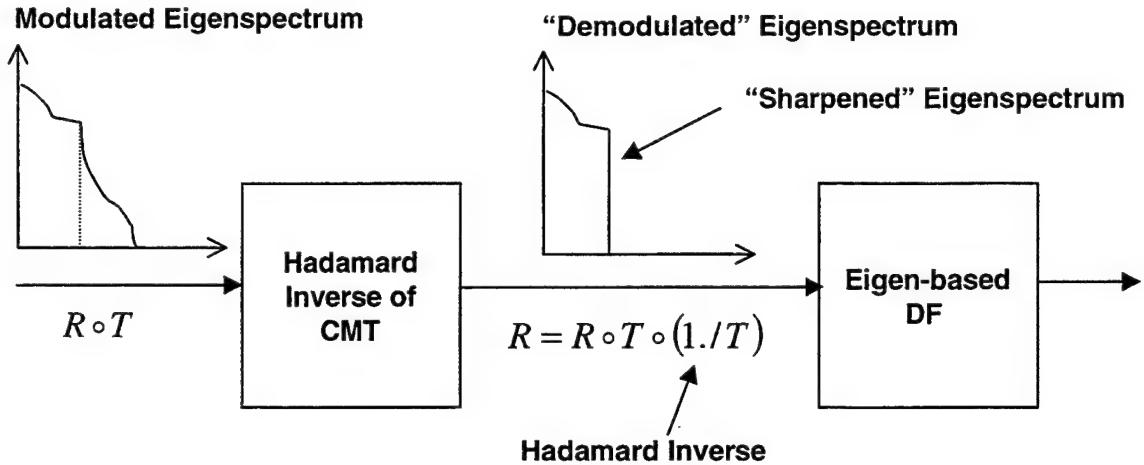
Figure 6 Illustration of the ICMT method for "demodulating" the blurred eigenspectrum prior to eigenbased direction finding.

[9] J. R. Guerci, J. S. Goldstein, I. S. Reed, H. Nguyen, P. M. Techau and J. S. Bergin, "Optimal reduced-rank STAP for circular adaptive arrays," *Proceedings of the Space-Time Adaptive Processing Methods for Circular Ring Arrays with Application to Navy Airborne Surveillance Radar,"* George Mason University, Fairfax, VA, Feb. 22-23 1999.

[10] M. A. Zatman, "Circular array STAP," *Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop*, March 10-11, MIT Lincoln Laboratory, Lexington, MA 1999.

[11] M. A. Zatman, "Production of adaptive array troughs by dispersion synthesis," *Electron. Lett.*, vol. 31, no. 25, pp. 2141-2142, 1995.

[12] R. J. Mailloux, "Covariance matrix augmentation to produce adaptive pattern troughs," *Electron. Lett.*, vol. 31, no. 10, pp. 771-772, 1995.

[13] J. R. Guerci, "Theory and application of covariance matrix tapers for robust adaptive beamforming," *IEEE Trans. on Signal Process.*, vol. 47, no. 4, April 1999.

[14] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, New York: Cambridge Univ. Press, 1991.

[15] J. S. Goldstein and I. S. Reed, "Theory of partially adaptive radar," *IEEE Trans. on Aerosp. Electron. Syst.*, vol. 33, pp. 1309-1325, October 1997.

[16] S. U. Pillai, *Array Signal Processing*, New York: Springer-Verlag, 1989.

11

SITE-SPECIFIC PERFORMANCE BOUNDS FOR INTERFERENCE MITIGATION IN AIRBORNE RADAR SYSTEMS

P. M. Techau

ISL, Inc.

8130 Boone Blvd.

Suite 500,

Vienna, VA 22182

pmt@isl-inc.com

J. R Guerci

SAIC

4001 N. Fairfax Dr.

Suite 475

Arlington, VA 22201

JGuerci@trg1.saic.com

T. H. Slocumb

SAIC

4001 N. Fairfax Dr.

Suite 475

Arlington, VA 22201

tslocumb@trg1.saic.com

L. J. Griffiths

George Mason Univ.

SITE

Mail Stop 4A3

Fairfax, VA 22030

Griffiths@gmu.edu

ABSTRACT

This paper addresses the problem of interference suppression in airborne radar systems. Both hot clutter that is induced by interfering sources and cold clutter that results from the radar transmitter are considered. A new method is presented that allows calculation of the space-time covariance matrix that will be observed under specific topographical conditions. The method incorporates phenomenology observed under site-specific conditions as well as system effects such as array geometry, receiver filtering, and system bandwidth. This approach differs significantly from the usual approach in which sampled-data analysis is used to estimate the covariance matrix. The new approach allows direct calculation of the degree to which space-time adaptive processing can be used to mitigate hot and cold clutter components for a specific heterogeneous terrain environment. From these results, one can infer performance bounds which are tighter and thus more meaningful than the thermal noise floor limit.

1. INTRODUCTION

Space-time adaptive processing (STAP) has emerged as a key enabling technology for future radars. As conceived in [1, 2], it provides coupled, two-dimensional (2D), spatial (array elements) and temporal (successive radar pulses) degrees-of-freedom (DoFs) to allow for the simultaneous suppression of monostatic radar clutter (cold clutter) as well as *direct-path* sidelobe jammers. Unfortunately, significant amounts of jammer energy generally will be scattered from the ground into the mainlobe and sidelobes of the system. Referred to herein as "hot" clutter (other terms include terrain-scattered interference (TSI) and jammer multipath), this interference cannot be mitigated using spatial techniques alone since it is present in the mainbeam. To jointly mitigate hot and cold clutter, both fast time (range delay) and slow time (PRI delay) temporal DoFs (as well as spatial DoFs) are needed. Joint processing of all three dimensions simultaneously is referred

to as three-dimensional (3D) STAP. While significant work in mitigating the effects of hot clutter has been accomplished including the assumption of factored approaches (for example, [3]- [5]), as well as 3D STAP [6], their performance has generally been gauged based on particular simulations involving sample statistics or simple clutter statistics models. The absence of 3D performance bounds makes it difficult to establish their absolute efficacy. However, establishing such bounds is complicated by generally heterogenous terrain, thereby all but precluding the adoption of conventional stationary clutter models.

In this paper, we present a method for establishing bounds on cold and hot clutter mitigation by using known scattering parameters. This results in a hybrid stochastic-deterministic covariance matrix formulation where the covariance matrices are formulated from deterministic (i.e. scatter amplitude based on digital terrain data) and random (i.e. jammer waveform) components. These data are available from site-specific digital terrain data [7, 8]. This method provides a performance bound that is tighter than the thermal noise floor limit usually given. We illustrate by example the application of the performance bound to a hypothetical system.

2. STOCHASTIC AND DETERMINISTIC COVARIANCE MATRICES

The need for STAP arises in airborne MTI (moving target indicator) radar due to the dynamic and generally simultaneous presence of jamming and clutter. Monostatic narrowband cold clutter arising from the reverberations of the transmit pulse from the surrounding terrain is spread in Doppler due to ownship motion and antenna sidelobes [2].

We designate $\mathbf{x}_{cc}(t, m) \in C^N$ to represent the cold clutter received by the N-element array from the mth pulse. These vectors are concatenated to provide the standard space (slow) time data vector [2] corresponding to the N element array output vectors ("snapshots") for a fixed range bin for each of the M pulse returns in

a coherent processing interval (CPI). We see from [7, 8] that a natural bound on performance is achieved using the weight vector

$$\mathbf{w}_o = \frac{\left(\mathbf{R}_{cc} + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{a}(\theta_d, f_d)}{\mathbf{a}^H(\theta_d, f_d)\left(\mathbf{R}_{cc} + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{a}(\theta_d, f_d)} \quad (1)$$

where,

$$\mathbf{R}_{cc} \triangleq \sum_{p=1}^{P_{cc}} |\alpha_p|^2 \, \mathbf{v}(\theta_p, f_p)\mathbf{v}^H(\theta_p, f_p) \quad (2)$$

and α_p is the complex amplitude of the scattered energy from the pth patch in the range bin, $\mathbf{v}(\theta_p, f_p)$ is the space (slow) time steering vector for the pth patch, $\sigma^2 \mathbf{I}$ is the covariance associated with the zero mean sensor noise, $\mathbf{a}(\theta_d, f_d)$ is the desired space time constraint vector, and P_{cc} is the total number of scatterers within the range bin of interest. This weight vector results from a hybrid *deterministic-stochastic L_2-norm* optimization performed in the angle-Doppler domain that assumes that the α_p are known and that sensor noise is the only random component of the optimization.

Since hot clutter signals consist of time-delayed and Doppler-shifted versions of the direct path signal scattered from the ground, we consider STAP using fast time taps. The data output from each element may be arranged as a complex vector. Likewise, the outputs of each level of taps may be arranged as a vector. We consider the output of the lth set of time taps. (There are L fast time taps in all.) Since the scattered signals are scaled, time-delayed versions of the jammer signal, this data vector may be represented by:

$$\mathbf{x}_{hc,l}(t) = \sum_{p=1}^{P_{hc}} \beta_p \mathbf{v}(\theta_p) s_j(t - \tau_p - l\tau) e^{j2\pi f_p(t - l\tau)}$$

where β_p is the complex magnitude of the signal scattered from patch p, $\mathbf{v}(\theta_p)$ is the array response vector (steering vector) to the signal with angle of arrival (AoA) θ_p, τ_p is the bistatic time delay from scattering patch p, $s_j(t)$ is the jammer waveform, P_{hc} is the total number of hot clutter scatterers, τ is the delay between time taps (e.g. the sample interval), and the direct path signal has been included in the summation. Each of the terms in the above equations is assumed to be known (deterministic) with the exception of the specific realization of $s_j(t)$ which is assumed to be a zero-mean Gaussian random process with known power spectral density (PSD). The precise treatment of the β_p is described in [7]. The covariance matrix is simply the expectation of the outer product of the concatenated data vector with itself. The cross covariance between

the ith and k^{th} fast time tap outputs is:

$$\begin{aligned} \mathbf{R}_{ik}(t) &= E\left[\mathbf{x}_{hc,i}(t)\mathbf{x}_{hc,k}^H(t)\right] \\ &= \sum_{p=1}^{P_{hc}} \beta_p \mathbf{v}(\theta_p) \sum_{q=1}^{P_{hc}} \beta_q^* \mathbf{v}^H(\theta_q) \times \quad (3) \\ &\quad r_j(\tau_p - \tau_q + (i - k)\tau) \times \\ &\quad e^{-j2\pi(if_p - kf_q)\tau} e^{j2\pi(f_p - f_q)t} \end{aligned}$$

where $r_j(\tau) = E[s_j(t)s_j^*(t - \tau)]$ is the jammer correlation function and we note that since each submatrix is not a function of only the difference between i and k, the resulting covariance matrix is not block Toeplitz due to the presence of platform motion-induced Doppler shifts. Note that in this case a double summation (as compared to the cold clutter formulation) has resulted. This stems primarily from the correlations which result from the simultaneous presence of all scattered components in the signal.

3. 3D STAP BOUND

As seen in [7, 8], a bound on joint hot and cold clutter mitigation may be established using 3D STAP and a hybrid stochastic-deterministic L_2-norm optimization employing the combination of the cold clutter, hot clutter, and noise covariance matrices. The application of the weight vector

$$\mathbf{w}_o(t) = \frac{\mathbf{R}_{\mathbf{xx}}^{-1}(t)\mathbf{a}_{3D}(\theta_d, f_d)}{\mathbf{a}_{3D}^H(\theta_d, f_d)\mathbf{R}_{\mathbf{xx}}^{-1}(t)\mathbf{a}_{3D}(\theta_d, f_d)} \in C^{LMN} \quad (4)$$

where

$$\mathbf{R}_{\mathbf{xx}}(t) = E\left[\mathbf{x}_{3D}(t)\mathbf{x}_{3D}^H(t)\right] \in C^{LMN \times LMN}, \quad (5)$$

$$\mathbf{a}_{3D}(\theta_d, f_d) = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}^H \otimes \mathbf{a}(\theta_d, f_d),$$

and $\mathbf{x}_{3D}(t)$ is the 3D STAP data vector described in [7, 8] results in a bound on signal-to-interference plus noise ratio (SINR) performance. The expectation in (5) is over the noise and unknown (but known PSD) jammer waveform.

For the determination of the performance bound, we consider only the clutter from the range bin of interest. Also we assume that the radar PRI is long compared to the inverse of the system bandwidth. Finally, we assume that there is no oversampling. This results in the interference covariances having the structure shown in [7].

4. PERFORMANCE SIMULATIONS

The phenomenology modeling tool used for this analysis is the Splatter, Clutter, and Target Signal (SCATS) model [7, 8]. Since this model provides for the complete

parameter	value
bandwidth (B)	100 kHz
no. element (N)	8
no. pulses (M)	32
fast time taps (L)	5
radar velocity	125 m/s 0° az
jammer velocity	165 m/s 0° az
jammer correlation	$\text{sinc}(\pi B\tau)$
sample rate	B
clutter-to-noise ratio	60 dB/element and pulse
jammer-to-noise ratio	50 dB/element
pulse repetition freq.	725 Hz

Table 1: System parameters used in simulations.

characterization of the signal environment produced by a transmitter (i.e. a jammer or a radar transmitter) in terms of the signal strength, delay, Doppler, and angle of arrival (AoA) of each scattered signal, we may use this model to find the 3D STAP covariance matrices as shown in [7].

The simulation scenario is shown in Figures 1 and 2 with the platform locations and heights shown (the height above local terrain is shown along with the height of the local terrain in parentheses). The system parameters used in the simulation are shown in Table 1. For this analysis we ignore the effects of clutter ambiguities.
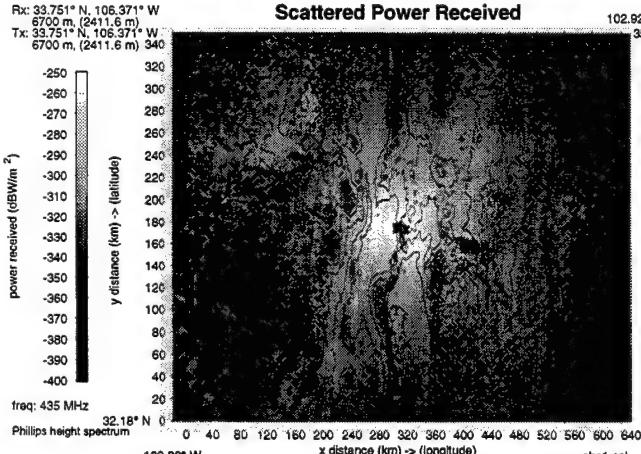
Figure 1: Cold clutter received on a single element for system located over North Oscura Peak.

Figure 1 shows a plan view of the cold clutter returns for the radar system. The map shows scattered power received per unit area (in this case dBW/m^2) assuming unity transmit power. These results are then scaled as desired in the adaptive analysis. Figure 2 shows a similar plot of the received hot clutter energy.
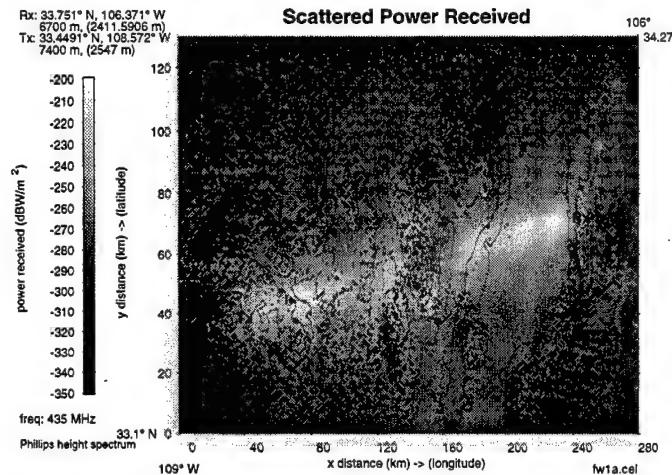
Figure 2: Hot clutter returns received on a single element from an airborne jammer for the system located over North Oscura Peak.

Note the strong returns along the line between the jammer and the receiver. This is often referred to as the glistening region and results from the near-specular geometry for the scattering cells within this region.

The measure of performance we considered in this analysis is SINR loss [2]. SINR loss is a metric that compares the beamformer output SINR to the SNR in an interference-free environment. Figure 3 shows results for the cold clutter and jammer-induced interference separately. Shown are a single-bin post-Doppler clutter result and the loss resulting from the use of spatial adaptive processing on the hot clutter. A 60 dB Chebychev Doppler taper is used throughout this analysis. The clutter range bin is 50 km, as indicated in the figure, while the Doppler bin of interest for all cases is 181.25 Hz (Doppler bin 8). We see the effects of using a site-specific model for the interference. For the cold clutter, we have additional losses at approximately azimuth 230° while the hot clutter shows additional losses due to the hot clutter at azimuths well away from the jammer.

We now show the 3D STAP performance bound for joint hot and cold clutter mitigation. The range bin and Doppler filter under consideration remain unchanged, as well as the assumed interference levels. Figure 4 shows two sets of plots. The SINR loss curves are shown for cold clutter plus noise, hot clutter plus noise, and hot and cold clutter plus noise. The corresponding eigenvalues are also shown. The SINR loss curves have the expected notches corresponding to the jammer and clutter azimuths. We see that the expected number of significant clutter-only eigenvalues (39, according to Brennan's rule) has been multiplied by the
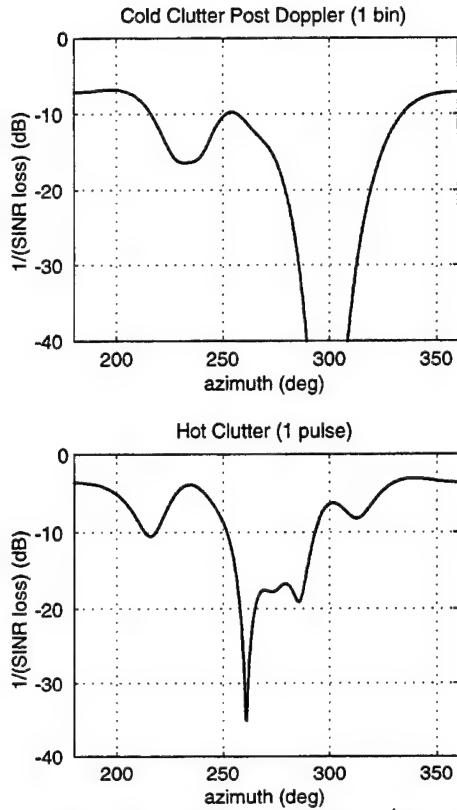
Figure 3: Top: cold clutter plus noise (post-Doppler 1bin) SINR losses. Bottom: hot clutter plus noise SINR loss (spatial adaptation on the instantaneous covariance.

Figure 4: Top: 3D STAP performance bound SINR loss curves for cold clutter plus noise, hot clutter plus noise, and hot and cold clutter plus noise. Bottom: eigenvalues correponding to the covariance matrices used to calculate the 3D performance bound.

number of fast time taps. This is expected since there is no range bin-to-range bin correlation of the clutter. Likewise the number of large hot clutter eigenvalues corresponds to the number of fast time taps and pulses. This results from the lack of pulse-to-pulse correlation of the hot clutter as well as the lack of correlation of the jammer direct path between fast time taps. The SINR loss curves shown here represent the best possible performance achievable in the simulated interference environment for the given system parameters.

We now consider the analysis of several post-Doppler interference mitigation architectures. These results are not intended to advocate any particular mitigation architecture but rather to illustrate the application of the performance bound as well as the theoretical framework for assessing and isolating the contribution of various factors to the losses observed. First we consider the decorrelating effects of Doppler on the hot clutter by considering the covariance matrix that results from the ensemble average of sample covariance matrices formed from the averaging of samples over a single PRI [7]. This result is shown in Figure 5, where we

have compared the bound to the result assuming training over a single PRI. Note that since this result is from the ensemble average over all possible sample covariance matrices, the additional losses incurred relative to the bound are attributable only to the decorrelating effects of Doppler, and not due to finite sample statistics.

The effects of cold and hot clutter both individually and jointly are shown in Figure 6. Here we consider three post Doppler architectures: 1 bin/1 fast time tap, 1 bin/5 fast time taps, and 3 bin/5 fast time taps. We see that the addition of fast time taps does not affect the cold clutter mitigation as expected, due to the lack of correlation of the cold clutter between the fast time taps. In the case of the hot clutter, the fast time taps improve performance but only a very slight amount. This is due to the hot clutter being thoroughly decorrelated by the Doppler processing. When multiple Doppler bins are used, both cold and hot clutter mitigation are improved. This is due to the usual expected narrowing of the clutter notch for multi-bin STAP and

16
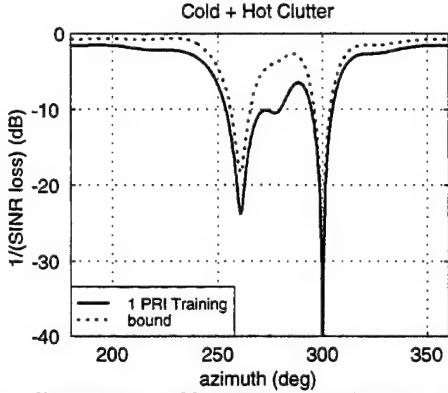
Figure 5: Illustration of losses incurred due to Doppler decorrelation.

Figure 6: Results for seperate and joint hot and cold clutter only mitigation using several post Doppler mitigation architectures.

the Doppler compensation on the hot clutter which results from the use of several Doppler bins. The same trends hold for the joint hot and cold clutter mitigation. In each case the 3D STAP performance bound is shown for each type of interference. Note that for each individual type of interference, this corresponds to the appropriate 2D STAP bound (space slow time for cold clutter, and space fast time for hot clutter).

In Figure 7 we show the results for varying the number of Doppler bins used in the processing for one and five fast time taps. These results are compared to the one tap and five tap performance bounds, respectively. We see that the inclusion of Doppler DoFs in the processing greatly improves performance and that when nine of the 32 Doppler bins are used, we are within several dB of the bound in the five tap case, slightly poorer performance (relative to the one tap bound) results for the one tap case.

We now consider a case of five Doppler bins and five fast time taps. In Figure 8, we show the effects of training over a single PRI. This result is again from an ensemble average over all possible sample covariance matrices, thus we are seeing only the effects of Doppler decorrelation.

Finally, we consider a result which illustrates tradeoffs of varying types of DoFs. In Figure 9, we show the comparison between the one fast time tap performance bound (bnd (1t), 256 DoFs) and a five fast time tap/five Doppler bin result (200 DoFs). Here we observe that with fewer DoFs, the mitigation architecture which combines temporal and Doppler DoFs achieves comparable performance to the architecture which uses only slow time taps (or Doppler) DoFs.

5. SUMMARY

In this paper, we have illustrated the application of a new 3D STAP performance bound on the joint mitigation of hot and cold radar clutter when site-specific digital terrain data is used. This approach has been illustrated with a scenario related to the DARPA Mountain Top experiments [9]. We have observed that fast time taps are ineffective in post Doppler processing of combined hot and cold clutter interference without simultaneously using Doppler DoFs. In addition, we have observed losses that result from Doppler decorrelation when estimating the covariance matrix with sample averages. This has illustrated how this method allows effects such a algorithm factorization and Doppler decorrelation to be isolated in the performance analysis.
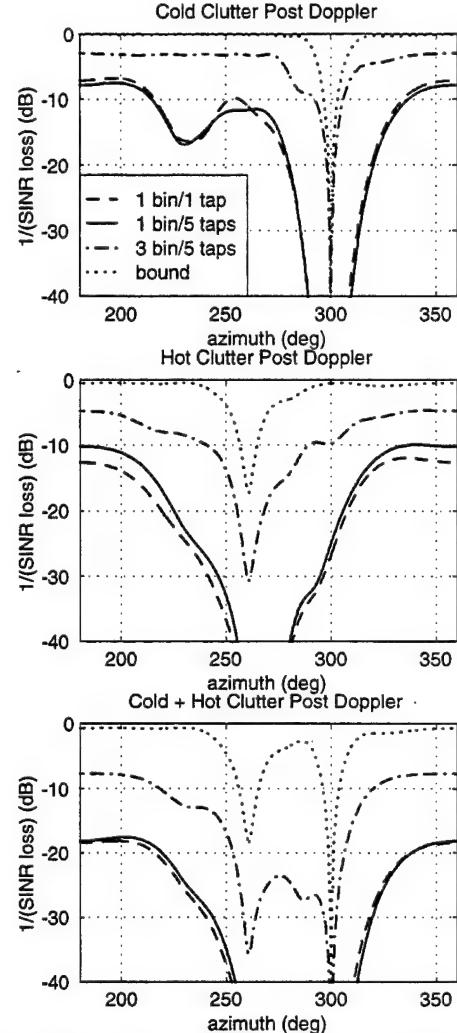
17

Post Doppler Results 1 Fast Time Tap

Post Doppler Results 5 Fast Time Taps
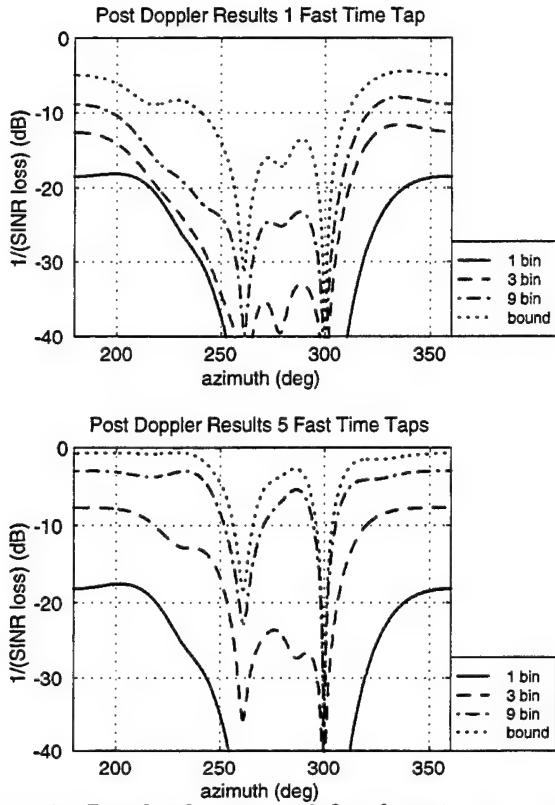
Figure 7: Results for one and five fast time taps and varying numbers of Doppler bins.

Post Doppler Results (5 bin/5 tap)

Figure 8: Doppler decorrelation effects due to finite training interval on a post Doppler architecture.

Post Doppler Results (5 bin/5 tap)

Figure 9: Comparison of one tap bound and a five bin/five fast time tap post Doppler architecture.

6. REFERENCES

[1] L. E. Brennan and I. S. Reed, "Theory of adaptive radar," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 9, March, 1973.

[2] J. Ward, "Space-time adaptive processing for airborne radar," Technical Report 1015, MIT Lincoln Laboratory, Lexington, MA, December 1994.

[3] T. W. Miller and J. M Ortiz, "An overview of issues in hot clutter mitigation," Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop, March 13-15, 1996, MIT Lincoln Laboratory, Project Report ASAP-4, Lexington, MA.

[4] L. J. Griffiths, "Adaptive hot clutter mitigation using multiple linear constraints," Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop, March 13-15, 1996, MIT Lincoln Laboratory, Project Report ASAP-4, Lexington, MA.

[5] D. F. Marshall and R. A. Gabel, "Simultaneous mitigation of multipath jamming and ground clutter," Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop, March 13-15, 1996, MIT Lincoln Laboratory, Project Report ASAP-4, Lexington, MA.

[6] R. L. Fante and J. A. Torres, "Cancellation of diffuse jammer multipath by an airborne adaptive radar," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 31, April, 1995.

[7] P. M. Techau, J. R. Guerci, T. H. Slocumb, and L. J. Griffiths, "Performance bounds for hot and cold clutter mitigation in airborne radar systems," in review with *IEEE Transactions on Aerospace and Electronic Systems*.

[8] P. M. Techau, J. R. Guerci, T. H. Slocumb, and L. J. Griffiths, "Performance bound for interference mitigation in radar systems," *Proceedings of the 1999 IEEE Radar Conference*, Waltham, MA, April 20-22, 1999.

[9] G. W. Titi, "An overview of the ARPA/Navy Mountaintop Program, *Proceedings of the IEEE Adaptive Antenna Systems Symposium*, November 1994.

Characterization of LMS and DMR Beamformers in the Presence of Loud Periodic Interferers

Geoffrey C. Street
Lockheed Martin Corp.
Manassas, VA, 20110-4157

Dr. Kristine Bell
George Mason University
Fairfax, VA 22030-4444

geoff.street@lmco.com, kbell@gmu.edu

ABSTRACT

Sonar system adaptive array processing is designed for slowly varying noise fields. The rapid expansion of deep-water offshore oil exploration has introduced loud impulsive periodic interferers into these systems. Standard LMS adaptive beamforming algorithms do not effectively eliminate these types of sources. In this paper we propose and examine the performance of several adaptive beamforming techniques designed to mitigate both periodic and static interference.

1. INTRODUCTION

Adaptive Least Mean-Square (LMS) beamformers are efficient algorithms, providing near optimum performance in slowly varying anisotropic noise fields [1]. The algorithm's adaptation step size is adjusted to match the dynamics of the noise and interference environment. Smaller step sizes provide improved asymptotic performance against a stationary noise field, but performance is penalized if the resulting time constant is longer than that of the noise field's actual temporal dynamics.

The rapid expansion of deep-water offshore oil exploration has introduced loud impulsive periodic interferers into the ocean's acoustic environment. Seismic profiling employs air or water guns to generate broadband pulses, strong enough to maintain adequate SNR for returns from the rock strata beneath the ocean floor. Repetition rate is dependant on the survey depth, varying from 10 to 80 seconds. LMS beamformers are slow to form nulls on these interferers because adaptation only occurs when the interferer is present in the data. Even after convergence, null depth is related to the average interference level, rather than the maximum level, and nulls are often very shallow.

In this paper we propose and examine the performance of several adaptive beamforming techniques designed to mitigate both periodic and static interference. The

algorithms are variations on robust LMS [2] and Dominant Mode Rejection (DMR) [3]-[5] techniques. When the advent of a periodic interferer is detected, an estimate of its direction of arrival (DOA) or array response vector is made. The first variation imposes first and second order null constraints in the direction of the interference. The second variation augments the array data vector with steering vectors pointed at each of the interferers. These steering vectors are spread in azimuth to account for DOA estimation inaccuracy. These algorithms have reasonable computational requirements and are robust to perturbations in the environment and array. Sets of simulations have been completed assessing performance in an idealized narrow band multi-interferer environment. The performance of each of the techniques in terms of interference suppression and computational complexity will be presented.

2. LMS PERFORMANCE

This section will review the configuration of the LMS beamformer used in this assessment, enhancement approaches, and results of simulation evaluations.

2.1 LMS Algorithm Overview

The baseline LMS beamformer used in this assessment is similar to that discussed in [1]. It utilizes linear constraints: a mainbeam distortionless constraint, and includes two beam shape constraints (at the 6 dB down points). In addition, it has a quadratic constraint to control the white noise gain, employing the scaled projection technique introduced in [2]. It is implemented as a generalized sidelobe canceler. The LMS beamformer was modeled as one of 'M_ω' narrowband beamformers. Only a single narrow band channel was configured for the following simulations. The input data to this channel was modeled as a complex sequence of FFT single bin outputs (referred to as a snapshot), one for each array element.

The basic algorithm is as follows:

$$w_q = C(C^H C)^{-1} f \qquad (1)$$
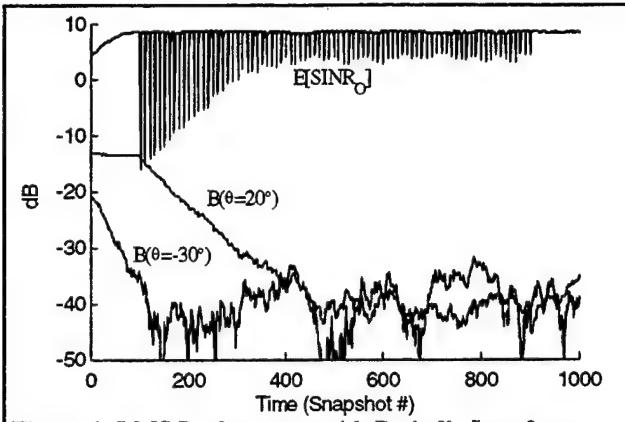
Figure 1 LMS Performance with Periodic Interference
LMS never eliminates the periodic interference. $SINR_O$ for a broadside 0dB SNR_I signal is shown, averaged over 20 independent trials. The broadside beam's response in the directions of the 30dB INR periodic source (20°) and the 10dB continuous source (-30°) illustrate LMS adaptation behavior.

Figure 2 $E[SINR_O]$ Performance Depends on INR
Increasing INR of other continuous sources lengthens convergence time for the periodic source, but improves asymptotic performance.

$$w_a(t) = w_a(t\text{-}1) + \alpha B^H x(t)[x(t)^H w_q - x(t)^H B w_a(t\text{-}1)] \quad (2)$$

$$\alpha = \gamma / (x(t)^H BB^H x(t)) \quad (3)$$

If $w_a(t)^H w_a(t) \geq 2w_q^H w_q$, then $w_a(t)$ is scaled so that the equality is satisfied. The final weight vector is:

$$w_o(t) = w_q - B w_a(t) \quad (4)$$

Where:

    C - Constraint matrix
    f - Vector of constraint levels
    B - null space of C, i.e. $B^H C = [0]$, $B^H B = I$
    $x(t)$ - Nx1 vector of array data at snapshot t
    $w_o(t)$ - Nx1 weight vector at sample snapshot t
    γ - LMS adaptive step size control
    ^{H} - Complex Conjugate Transpose

2.2 Seismic Profiler Model

Seismic profilers usually tow a horizontal array of air guns, up to 18 in each array. The air gun pulses are timed to focus most of the energy toward the ocean floor. Peak source levels of seismic profilers can be 260 dB re 1μPa, over a 400 Hz band. Actual reverberation includes both local medium non-linearities, as well as reflections from the ocean bottom and surface.

The simplified model used in this evaluation assumes the received signal is a single planewave. It is a sequence of pulses, each followed by a simple exponential decay, with a 1 second time constant. A 10 second pulse repetition interval was used for the simulations presented in this paper. A non-overlapped FFT output (snapshot) is completed every second. Each sequential snapshot was modeled as independent white Gaussian noise at a power level dictated by the exponential decay rate.

2.3 Impact to LMS performance

The performance of a standard LMS algorithm in an environment with a periodic interferer was simulated to determine its behavior under ideal conditions. This simulation used a 10 element line array with ½ wavelength spacing. A 0 dB SNR_I (element level) signal was injected at the broadside beam (0°). White Gaussian noise at 0dB was injected at each array element. A 30 dB single element Interference-to-Noise Ratio (INR) periodic interferer (10 sec interval) was injected at 20° from broadside. A 10 dB INR continuous interferer was injected at -30°. The beamformer was configured with a step size of $\gamma = 0.05$, and a 3 dB white noise gain constraint (however, this scenario was such that the unconstrained weight vector always satisfied the white noise gain constraint).

Figure 1 shows how an LMS beamformer behaves in this environment. In the period from t=1-100, the beamformer successfully adapts a null at the continuous interferer ($\theta = -30°$), achieving a 9dB output Signal to Interference plus Noise Ratio ($SINR_O$) for the desired signal at broadside (smaller step sizes would allow the asymptotic $SINR_O$ to approach the ideal 10 dB white noise gain). Since this source dominates during this period, the beamformer converges with a time constant of $1/\gamma$.

The periodic interferer starts at t = 100. Its immediate impact to $SINR_O$ results from the -13dB sidelobe level at t=100. The standard LMS adapts a null at this direction at a slower rate than for a continuous source. Its asymptotic null depth allows this interferer's pulse peaks to degrade the desired signal's $SINR_O$. The convergence

rate and depth of the null steered at the periodic source is a function of the two sources' INRs, the average power of the periodic source over the time constant of the beamformer, and the LMS step size γ. In this example, the periodic source's average power is 8dB below it's peak power.

As the LMS step size parameter is increased, the convergence rate improves, with the penalty of poorer asymptotic performance against all types of sources. A key performance trade in this environment is matching the step size to the dynamics of the sources. Step sizes on the same order as the inverse of the interferer's repetition interval seem to provide the best convergence rates. Larger step sizes give better asymptotic performance.

Another issue is the dependence of the asymptotic convergence on the continuous source's INR. Figure 2 shows the prior scenario with the source at -30° having an INR of 20 and -10 dB. Higher continuous source INR improves the converged null depth for the periodic source. Lower levels result in poorer converged null depth This is a natural result of the LMS algorithm's well documented convergence rate characteristics [6].

LMS adaptation rate is a function of the spread in the levels of the array's cross spectral density matrix's (CSM) eigenvalues. The convergence rate for each eigenvalue and eigenvector pair (λ_i, U_i) is proportional to the ratio of $\lambda_i / \lambda_{max}$, i.e. the smaller the ratio, the slower to converge a null at U_i. With no other source impinging on the array in between pulses, the LMS ABF can 'rapidly' re-optimize $SINR_O$ to the ambient noise field. This will generally degrade the developing null in the direction of the (now absent) periodic source. If a loud continuous source exists, its eigenvalue/vector dominates during the inter-pulse interval. Thus with a continuous interferer, the beamformer is slow to optimize output power for the background noise in the absence of the periodic source. As a result, the developing null at the DOA of the periodic source sees reduced inter-pulse degradation.

2.4 Approaches to Improve Performance

Once profiling commences, it typically continues for long periods of time, with fairly constant operating parameters. Tow speeds are less than 5 knots, and the dynamics are relatively stable. The array receive level for each pulse is loud, with 20-30dB/√Hz INRs not unusual over a 400 Hz bandwidth. Signals of this nature support accurate direction of arrival (DOA) estimates. Because of the stable dynamics, this estimate is a good predictor of the next pulse's DOA.

This suggests approaches that should improve beamformer performance over standard LMS implementations. Two were considered. The first approach forces a broad null at the estimated direction of the interferer. The second approach augments the data vector with steering vectors pointed at the interferer.

Both approaches require a method to detect the advent of profiling. The method initially used in this study [7] compared the current snapshot's array power level against an exponentially weighted average:

$$\hat{\sigma}_{PD}^2(t) = (1 - \mu_D) \sum_{k=1}^{t} \mu_D^{t-k} x(k)^H x(k) \qquad (5)$$

where μ_D is usually between 0.9 and 0.95. If the power received at the array for any one snapshot exceeded this statistic, then the advent of profiling was assumed. A similar statistic with a longer time constant was calculated to determine when profiling activity ceased. This method is effective when only a single profiler is operating. In many situations, there several profilers concurrently operating in any one region. When there are, it is advantageous to determine the advent or cessation of each individual source's activity. The above approach does not easily support this need.

To address this scenario, we took advantage of the fact that applications of interest form a complete azimuthal beam set. A detection and loss statistic, similar to (5), was created for each beam. Once the advent of profiling was detected, its DOA was estimated. A simple beamscan [1] algorithm was selected to estimate DOA. This algorithm has the advantage of low computational requirements. These high INRs does not require a more complex DOA estimation technique. The beamscan was limited to the spatial region of the detecting beam's null-to-null width.

Three array steering vectors were then calculated based on this DOA estimate:

$$v_{I, I+\Delta, I-\Delta} = e^{\left(j\frac{2\pi}{\lambda} \underline{d} \cos(\theta_I \pm \theta_\Delta) \right)} \qquad (6)$$

where $\underline{d}$ are the element positions, θ_I is the interference DOA estimate, and θ_Δ is selected based on the uncertainty of the DOA estimate.

2.5 Forced Null Constraints

In this approach, the constraint matrix and levels, C and f in (1), were augmented with $[v_I, v_{I+\Delta}, v_{I-\Delta}]$ to force a null at these arrival directions (a 2nd order derivative null could also be used). This was not done for the beams
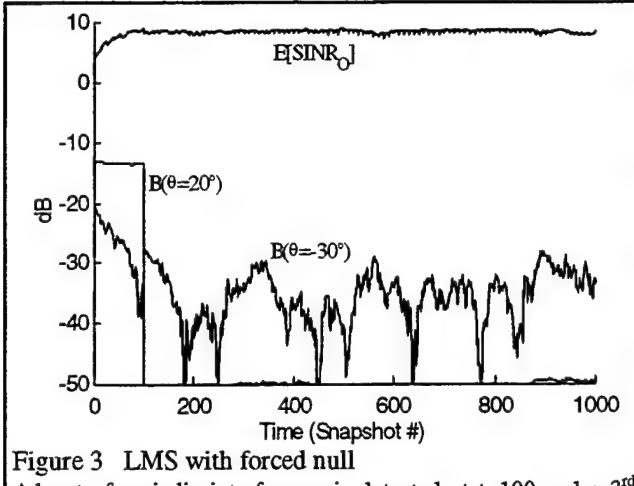
Figure 3   LMS with forced null
Advent of periodic interference is detected at t=100 and a 3rd order null forced. DOA estimation error allows some leakage of the interferer. Note the degradation of the null at θ=-30° when the null is forced.

Figure 4   LMS with Augmented Data Vector
Advent of periodic interference is detected at t=100 and the data vector is augmented with three steering vectors closely spaced about the estimated DOA. Note the degradation of the null at θ=-30° as the beamformer adapts a null to the augmented data matrix.

within ½ beamwidth, null-to-null, of the interference DOA. In order to maintain the adapted nulls at other interferers, the adapted weight vector, w_a, was recalculated as follows:

$$\tilde{w}_a(t-1) = -\tilde{B}^H w_o(t-1) \qquad (7)$$

$\tilde{B}$ is the null space of the updated constraint matrix. The LMS weights are then updated (2)-(4) with the current snapshot.

Simulations were run over a variety of different scenarios to evaluate this approach. The broad null is effective at nulling out the periodic interferer. Only one performance issue was noted with this approach. In some cases the existing nulls pointed at other interferers deteriorated when (7) was applied. This depended on the relative positions of these interferers to the location of the forced null(s).

Figure 3 shows the results of one of these cases. Plotted are the E[SINR_O] for a single beam, along with beamformer response in the direction of each interferer. It shows the effective nulling of the periodic interferer, with the null depth generally below -50dB. With DOA estimation errors, the null is not perfect and Figure 3 shows the beamformer response at 20° occasionally increasing slightly above -50dB.

Figure 3 also shows deterioration of the null at -30° when the null is forced (t=100). This happens when the adapted weight vector, $w_a(100)$, is projected onto the new null space of the updated constraint matrix (7). The LMS algorithm does quickly re-adapt to the pre-existing nulls. Since the adaptation loop operates in the null space of the

constraint matrix, the periodic interferer's power does not affect the adaptation rate of the beamformer.

2.6   Augmented Array Data Vector

In this case, the snapshot data vector, $x(t)$ used in (2), was augmented with the scaled sum of the array steering vectors in (6):

$$\tilde{x}(t) = x(t) + \hat{\sigma}_I \left(\rho_1 v_I + \rho_2 v_{I+\Delta} + \rho_3 v_{I-\Delta} \right) \qquad (8)$$

where ρ_1, ρ_2, and ρ_3 are complex zero mean independent Gaussian random variables, and the scaling factor $\hat{\sigma}_I$ is the square root of the peak element power. The augmented array data vector, $\tilde{x}(t)$, is then used in (2) and (3). The LMS algorithm adapts nulls at these DOAs at a rate proportional to $1/\gamma$.

Figure 4 shows the results for the same example used in Figure 3. Plotted are the E[SINR_O] for the broadside beam, along with that beam's response in the direction of both interferers. It shows the beamformer adapting an effective null at the periodic interferer. While not as effective as forced nulls, the results should be very satisfactory for periodic interferers that operate nearly continuously. It does have the advantage of requiring significantly less computation than the forced null approach.

Note the degradation of the null at θ=-30° between 100≤t≤200 as the beamformer adapts a null to the augmented data matrix. This is followed by a very slow re-convergence. This slow re-convergence results from

22
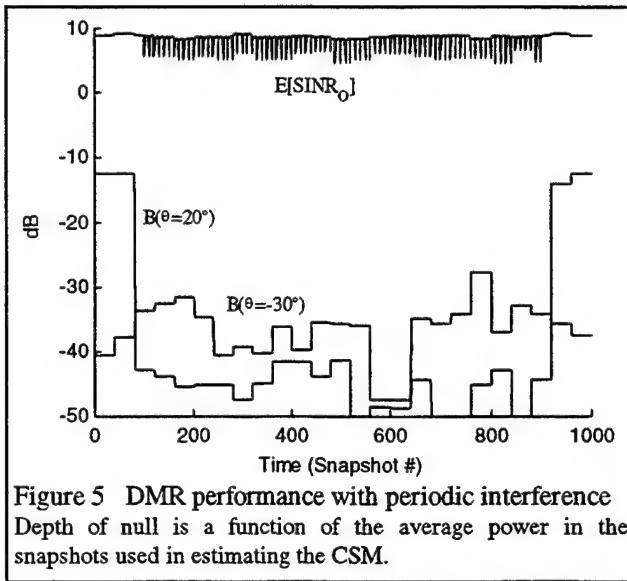
Figure 5   DMR performance with periodic interference
Depth of null is a function of the average power in the snapshots used in estimating the CSM.

Figure 6   DMR with Augmented Data Vector
Impact of periodic interference is virtually eliminated.

LMS convergence rate being proportional to the ratio $\lambda_i / \lambda_{max}$. For this example, the convergence rate is two orders of magnitude slower for the 10 dB source than the 30 dB augmentation.   This behavior argues for setting the augmentation to the minimum level that achieves satisfactory rejection of the periodic source.

2.7 LMS Summary

The forced null approach was the most effective approach. However, it   is also requires the most additional computations. If the number of constraints, N_C , is small relative to the number of array elements, N , this approach requires $O(M_\omega \cdot M_B \cdot N^2 \cdot N_C)$ additional operations when the constraints are updated (M_ω is the number of beamforming bands, M_B the number of beams, N the number of elements, and N_C the number of constraints). The constraints should be updated periodically, on the order of the LMS adaptation time constant.

Data matrix augmentation achieved nearly the same results after a few adaptation time constants. In typical applications, the adaptation time for the periodic source should not significantly impact system performance. This approach requires few additional computations, primarily the calculation of the array steering vectors, and additional computations should be on the order of $O(M_\omega \cdot N \cdot N_S)$ where N_S is the number of steering vectors. This will make it attractive for applications with limited computational resources.

If a particular application has continuous sources with fast temporal dynamics, forced null constraints may be preferable. Since it eliminates the loud source from the adaptive loop, the beamformer converges rapidly for these other sources.

3.   DOMINANT MODE REJECTION

Dominant Mode Rejection is an algorithm that provides fast convergence in a noise field with a large spread in it's CSM's eigenvalues. By only estimating the structure of the CSM's signal subspace, DMR can converge nulls in these directions at rate proportional to the size of the signal subspace. Since the size of the signal subspace is generally much smaller than the of number array elements, DMR requires fewer snapshots to achieve convergence than full eigenstate estimation approaches. Because it estimates the signal eigenstructure, DMR converges nulls at all signals at the same rate.   Therefore DMR should be an effective approach in scenarios such as those considered in this paper

3.1 DMR Algorithm Overview

The baseline algorithm used in this assessment is similar to that discussed in [3]. The D largest eigenvalues (and their eigenvectors) of the array's cross spectral density matrix, R , are determined:

Λ_D - DxD largest eigenvalue matrix

U_D - NxD eigenvector matrix

Then the CSM is estimated as:

$$\hat{R}_D = U_D \Lambda_D U_D^H + e \sigma_N^2 I_N \qquad (9)$$

were e is an empirically derived scalar enhancement factor and σ_N^2 is an estimate of the average N-D smallest eigenvalues. The inverse is simply calculated as:

23

$$\hat{R}_D^{-1} = \left[I_N - U_D \frac{\Lambda_D}{\Lambda_D + e\sigma_N^2 I_D} U_D^H \right] 1/(e\sigma_N^2) \qquad (10)$$

The DMR beamformer weights are then

$$\hat{w} = \frac{\hat{R}_D^{-1} v_s}{v_s^H \hat{R}_D^{-1} v_s} \qquad (11)$$

where v_s is the beam's array steering vector.

In this evaluation all eigenvectors and eigenvalues were calculated using singular value decomposition on sequential blocks of 40 array element snapshots. singular value decomposition was used to estimate the entire eigenstructure. The size of the signal subspace was estimated using the Akaiki Information Criterion [1]. We did not include data prior to the current block in calculating the eigenstate. The block size could be made larger in practice if the noise field temporal dynamics supported it.

A simulation was run using the same scenario described previously. The results are shown in Figure 5. DMR performance is similar to the LMS beamformer asymptotic performance described in the previous section. However, DMR estimates the complete CSM signal subspace and doesn't have the convergence issues associated with LMS. Because of this, the depth of the null directed at the periodic source is proportional to the average power over the data block, and is not affected by the presence of continuous sources. However, DMR $SINR_O$ performance is still degraded by the instantaneous power peaks.

3.2 DMR Profiler Enhancement

Again two approaches were considered for DMR: a forced null constraint and augmented data matrix. The primary benefit of the forced null constraint for LMS was immediate adaptation to the periodic source, and faster subsequent adaptation for the continuous sources. However, DMR has these features in its primary algorithm. Reducing the size of the signal subspace by 1 signal will only provide a modest improvement to convergence. Forcing null constraints completely eliminates the source, but at significant added computational cost. Therefore we explored augmenting the data vector.

This approach for DMR is virtually identical to the augmented data vector method described in section 2.6. The primary difference is that for DMR the input data is blocked, supporting a more accurate estimate of the sources' DOA and peak received power level.

Figure 6 shows the results of this enhancement. The periodic source's impact to $SINR_O$ is virtually eliminated. Further, since DMR estimates the full signal subspace, the nulls for the continuous sources are unaffected by the augmentation.

This approach can be implemented with only a small loading penalty, $O(M_\omega \cdot N \cdot N_S)$.

4. CONCLUSIONS

Periodic interferers were shown to have significant impact on the performance of both LMS and DMR beamformers. The primary cause is that both beamformers attempt to steer nulls that are proportional to the average power received from a given direction. The interferers of interest can easily have peak powers 10 to 30 dB above the general background noise. Three approaches were presented to mitigate the impact of periodic interferers, two for LMS and one for DMR beamformers. All three were effective at improving asymptotic performance.

A control approach was described. It takes advantage of a full azimuthal beamset to estimate the number of periodic sources and their DOAs. At the source levels where periodic sources are a problem, it provides a robust mechanism for injecting nulls or augmentation.

All evaluated approaches show promise in addressing this need. The next step is to test these approaches with actual single and multi-interferer sea data.

5. REFERENCES

[1] H. Van Trees, *Array Processing: Detection, and Estimation Theory, IV (draft)*, George Mason University, 1998.

[2] H. Cox, R. Zeskind, and M. Owen, "Robust Adaptive Beamforming", *IEEE Trans. Acoust., Speech, Signal Proc.*, ASSP-35, pp. 1365-1376, 1987.

[3] D. Abraham and N. Owsley, "Beamforming with Dominant Mode Rejection", Oceans '90 Conference Proceedings, pp. 470-475.

[4] H. Cox and R. Pitre, "Robust DMR and Multi-Rate Adaptive Beamforming", *31st Asilomar Conference on Signals, Systems, and Computers*, pp 920-924, Nov. 1997

[5] T. Redheendran and R. Gramann, "Initial Evaluation of the dominant mode rejection beamformer", *Jour. Acoust. Soc. Am.*, 104, pp.1331-1344, Sept 1998.

[6] B. Widrow and S.D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Englewood, NJ, 1985

[7] K. Bell and H. Van Trees, *Adaptive Beamforming for Interference Suppression*, Briefing Paper to Lockheed Martin, 1/8/1998.

EFFECTS OF ARRAY CONFIGURATION ON CIRCULAR ARRAY STAP PERFORMANCE

Jose H. Unpingco, Thomas W. Miller*

Raytheon Systems Company
Adaptive and Signal Processing
2000 East Imperial Highway
El Segundo, CA
jhunpingco@west.raytheon.com, twmiller@west.raytheon.com

ABSTRACT

In this paper, we consider how variation in the sector size of a circular array affects STAP performance for PRI-Staggered STAP. Sector size influences the achievable sidelobe level, and to a lesser extent, detection sensitivity and metric accuracy. As the sector size increases, the ability to control the sidelobes and cancel clutter/interference improves due to the increase in the number of degrees of freedom. In this paper, we study the effects of sector size with respect to PRI-Staggered STAP in terms of SINR, detectable velocities, clutter covariance eigenvalues, and usable Doppler space fraction. Additionally, we illustrate the problem of misalignment and training region size on STAP performance.

1. INTRODUCTION

The purpose of this article is to show how circular STAP (Space-Time Adaptive Processing) is influenced by sector size (the number of elements on a circular ring) with regard to PRI-Staggered STAP. To measure comparison, we consider the Signal to Interferor plus Noise Ratio Loss (SINRL), minimum detectable velocity (MDV), clutter covariance eigenvalues, and usable Doppler space fraction (UDSF)[1]. Increased sector size means using more elements on the circular array. For a linear array, the benefit of more sensors is straightforward: a sharper mainbeam that better localizes clutter. For a circular array, the benefit is not so obvious.

Also, we consider how training region size effects STAP perfomance and we illustrate why care must be taken when training the adaptive weights.

The outline of this article is as follows: First, we specify the scenario under consideration and discuss the element pattern and antenna patterns for different sector sizes. Next, we discuss and show the SINRL, MDV, clutter covariance eigenvalues, and UDSF for different sector sizes. Next, we illustrate an issue regarding training and STAP performance. Finally, some ideas regarding further research are proposed.

2. METHOD

Our method is to define a scenario, compute the SINRL, MDV, clutter covariance eigenvalues, and UDSF, and then see how these parameters change with sector size.

2.1 Scenario Parameters

We use the following scenario to study the effects of varying sector size.

- Sixty elements total in entire circular ring.

- Radius of ring = 12 feet

- PRF = 600 Hz

- Aircraft velocity = 100 m/s

- Altitude = 20,000 feet

- Clutter range ring = 75 km

- Transmit power at 75 km = 36 dB

- Uniform transmit taper

- Operating frequency = 425 MHz

- Wavelength = 70 cm

- 0.54 λ arc length inter-element separation

- Look direction is "boresight" (perpendicular to aircraft motion)

For PRI-Staggered STAP, we use a Doppler filter bank tapered with an 80 dB Chebyshev window. The overlap factor for PRI-Staggered STAP is K=3. For example, this means that for M=16 pulses, a 16-3+1=14 point discrete Fourier transform is computed on three overlapped data windows.

The same number of elements are used for both transmit and receive.

2.2 Array Patterns

Figure 1 shows the positions of half the elements in the ring (thirty elements). Sector size is symmetric with regard to the look direction (i.e. along the vertical line through x=0). For example, using four elements means using the two elements on either side of the vertical line through x=0 as in Figure 2. Each element has its own element pattern oriented outward. For example, Figure 3 shows the element pattern for the element on the ring at boresight. The element patterns are pointed away from the center of the circular ring.
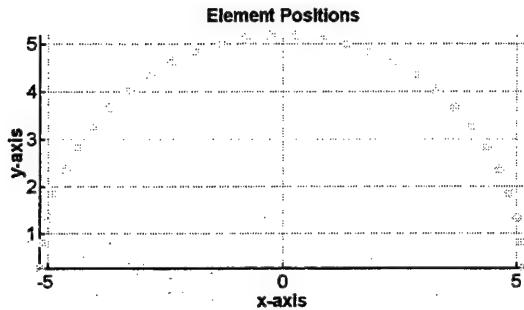
Figure 1 Thirty-element array configuration.

Figure 3 Antenna element pattern.

Figure 2 Four-element array configuration.
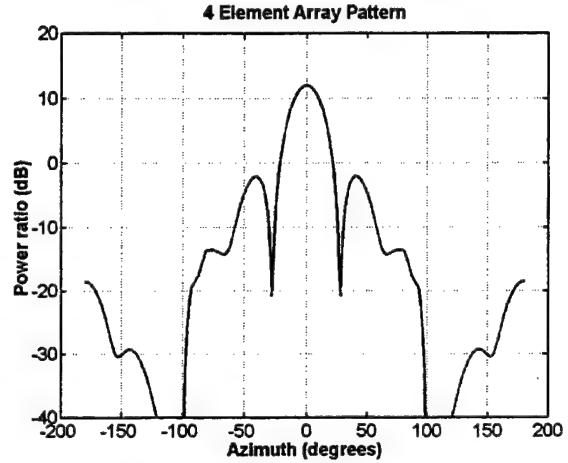
Figure 4 Four-element array pattern.

Figure 4, Figure 5, Figure 6, and Figure 7 show the array patterns for sector sizes four, ten, twenty, and thirty, respectively. The four-element array is most like a linear array, given the radius of the ring. Notice the peaks 30 dB down from the mainbeam at 180 degrees. This is due to the element pattern, which has a backlobe 30 dB down from its peak. The ten-element array shows less of this effect due to the increased curvature. However, the twenty and thirty-element array have large sidelobes at approximately 120 degrees due to more elements facing away from boresight. These large sidelobes announce the fundamental difference between the linear and circular array.

Figure 5 Ten-element array pattern.

26
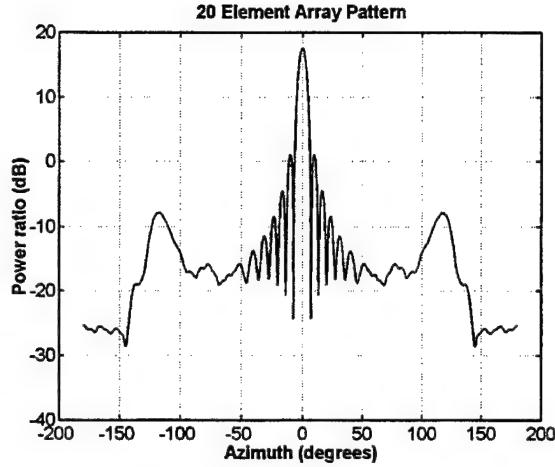
Figure 6 Twenty-element array pattern.
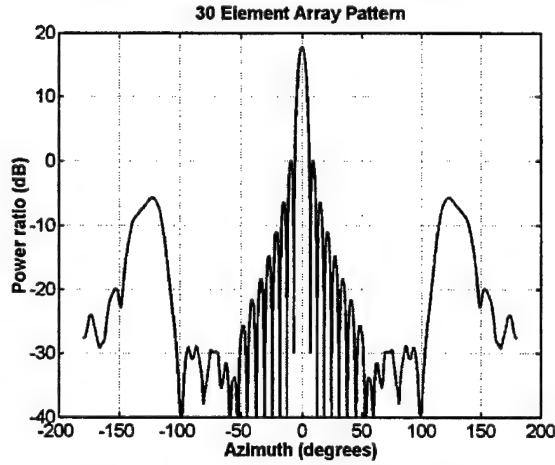
Figure 7 Thirty-element array pattern.

3. RESULTS

In this section, we consider how sector size influences each of our figures of merit: SINR Loss, MDV, clutter covariance eigenvalues, and USDF.

3.1 SINR Loss

Figure 8 shows the SINR loss for different sector sizes using M=16 pulses and PRI-Staggered STAP. The figure shows that increasing the sector size improves the SINR performance since the notch at boresight narrows. Note, however, that this improvement is asymptotic. That is, the improvement from N=20 to N=30 elements is proportional to the corresponding improvement between the N=4 and N=10 elements.
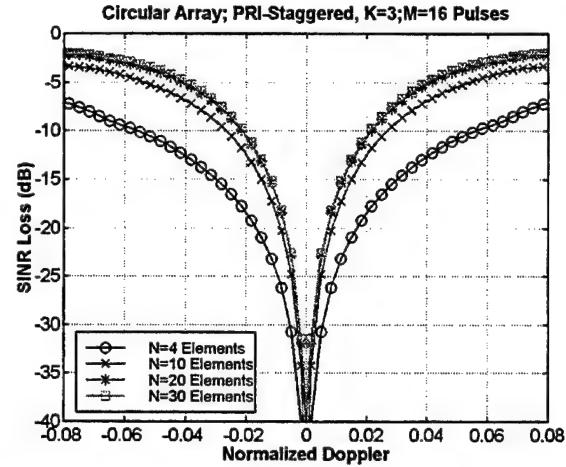
Figure 8 SINR Loss for PRI-Staggered (K=3) STAP.

3.2 Minimum Detectable Velocity

We measure the minimum detectable velocity by the width of the Doppler gap at –5 dB SINR Loss.
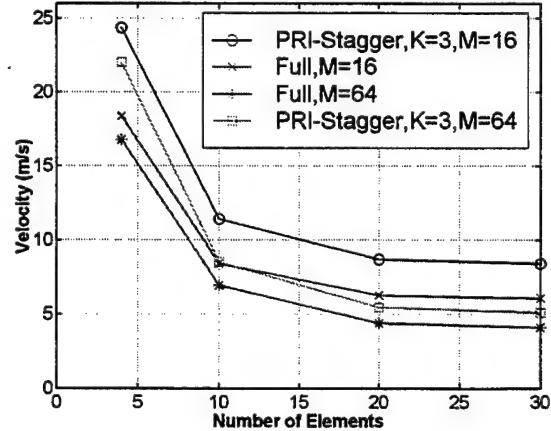
Figure 9 Minimum detectable velocity versus sector size.

In Figure 9, we have included the results for Full STAP with M=16 and M=64 pulses. Note that the MDV decreases asymptotically with increasing sector size. This means that the sharper mainbeam we obtain from using more elements on the circular ring has helped localize the clutter better. Also, note that the MDV for M=16 and M=64 pulse PRI-Staggered STAP do not intersect. In fact, the M=64 MDV result is a strict lower bound for the M=16 MDV result. This means that we have not observed a case where, given the same number of elements, it was possible to trade off the number of pulses for sector size.

Observe that the performance for Full STAP using M=16 pulses and the ten-element section closely matches the result for PRI-Staggered STAP with M=64 pulses. Finally, note that we once again see an asymptotic improvement

with growing sector size as we observed with the SINRL results earlier.

3.3 Clutter Covariance Eigenvalues

The eigenvalues of the clutter covariance matrix are shown in Figure 10 and Figure 11 for M=16 and M=64 pulses for varying sector sizes. The additive white Gaussian noise at the receivers has unit power (0 dB).

These figures show how many available dimensions exist in the data space which are orthogonal to the clutter. For example, the figure shows that for the thirty-element array with M=16 pulses, there are approximately fifty eigenvalues below the noise power out of the total ninety. This means that a signal that lies in the span of the corresponding eigenvectors of those fifty eigenvalues is immune from the clutter.

On the other hand, it also means that a signal with strong components in the span of the other forty eigenvectors will be adaptively nulled, since those components are indistinguishable from clutter.
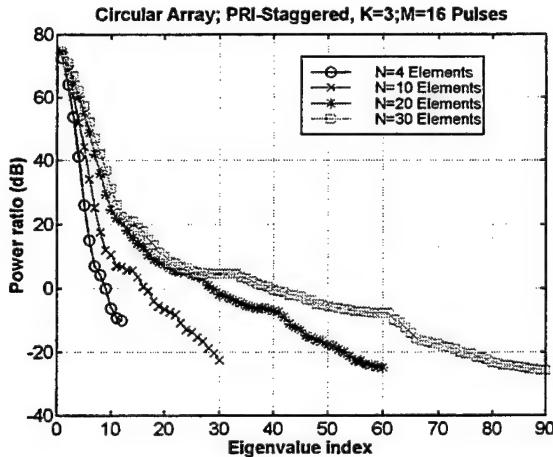
Figure 10 Clutter covariance eigenvalues for M=16 pulses.
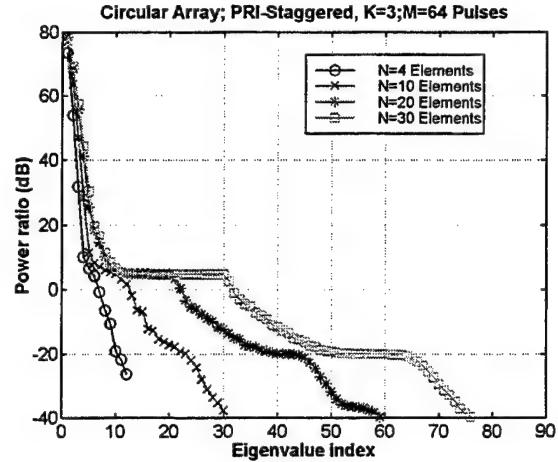
Figure 11 Clutter covariance eigenvalues for M=64 pulses.

An alert observation is that the eigenvalues do not level off at 0 dB as they would for Factored STAP and Full STAP. This is due to the overlapping of data in PRI-Staggered STAP.

Once again, we see the asymptotic improvement resulting from increased clutter size. This means that the improvement from N=4 to N=10 elements is not proportional to the corresponding improvement from N=20 to N=30 elements.

3.4 Usable Doppler Space Fraction

As the last measure of comparison between the sector sizes, we consider the Usable Doppler Space Fraction (UDSF). Figure 12 can be interpreted as the probability that the SINRL will be below –5 dB for a given number of elements (assuming all Doppler frequencies are equally likely).
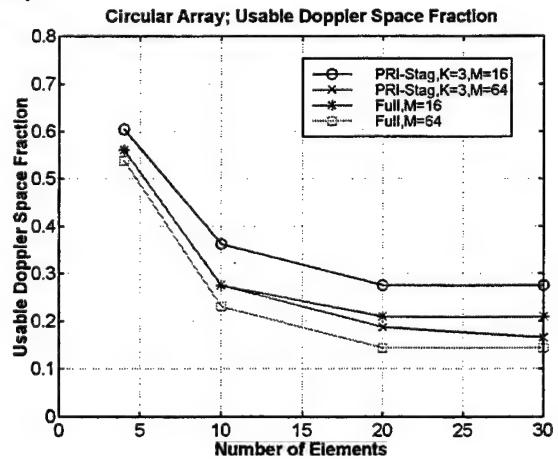
Figure 12 Usable Doppler space fraction versus sector size.

28

3.4 Training Region

Consider the situation where the beam is pointed off-boresight. This means that the mainbeam illuminates a curved section of the isodoppler contour. This causes a curve in the SINR Loss versus range for Full STAP as shown in Figure 13. For this example, the weights are trained at every clutter ring.
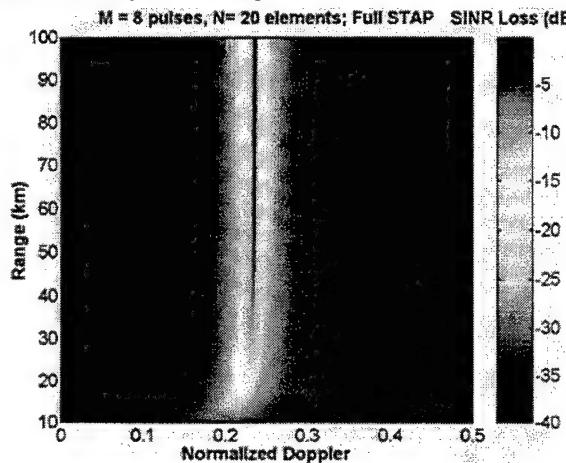
Figure 13 Look direction is 30 degrees off boresight. Weights trained on each clutter ring.

However, when the training region widens (e.g. 100-150 km), the SINR performance degrades to the extent that the clutter covariance varies with range. This means that the adaptive processing cannot "pick up" the near-field clutter since it is trained on far-field clutter as in Figure 14.
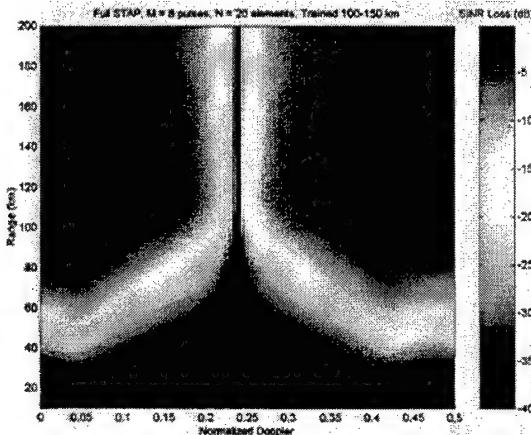
Figure 14 Weights trained on range region 100-150 km.

When we have a very wide training region (10-200 km), we lose the SINR curve we observed in Figure 15 and the near-field SINR notch widens.
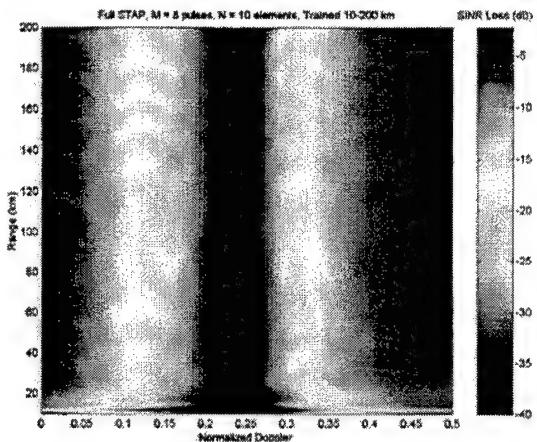
Figure 15 Weights trained on 10-200 km region.

Due to this effect, some care must be taken in the design of the training regions so as to not lose near-field performance for off-boresight look directions.

4.   CONCLUSIONS

The main conclusion is that using more sensors asymptotically improves STAP performance in terms of SINRL, MDV, and USDF. We observed a diminishing return when the sector size increased beyond twenty of sixty elements in the ring. We noted that PRI-Staggered (K=3) STAP was very near the performance of Full STAP in terms of all measures considered.

Some care must be taken in the design of the training regions for STAP. We illustrated a situation where the look direction was off-boresight and some SINR performance was lost due to the width of the training region.

The results presented here point to some potential further research: optimizing the training region size with regard to a specific STAP method, studying ways to taper the aperture and how tapers effect STAP performance, and studying how the clutter covariance varies with range and how this effects weight updating.

5.   REFERENCES

[1] J. Ward, "Space-Time Adaptive Processing for Airborne Radar," Technical Report 1015, MIT Lincoln Laboratory, Dec. 1994.
[2] R. Klemm, "Adaptive clutter suppression for airborne Phased Array Radars," IEE Proceedings, Vol. 130, Pts. F and H, No. 1, February, 1983, pp. 125-131.

Synthesis of Adaptive Monopulse Patterns

Ronald L. Fante

The MITRE Corporation
202 Burlington Road
Bedford, MA 01730-1420
rfante@mitre.org

ABSTRACT

A procedure is developed to simultaneously synthesize sum and difference patterns for Space-Time Adaptive Processing (STAP), in such a way that a specified monopulse slope after adaptation is achieved.

ANALYSIS

Space-Time Adaptive Processing (STAP) is an effective method used by airborne radars for adaptively cancelling clutter and jammers, while simultaneously detecting targets. However, while it is straight forward to form adapted sum (Σ) and difference (Δ) beams, the adapted monopulse pattern Δ/Σ may have a highly distorted slope, rendering it ineffective for angular location [1]. In this letter we present an approach to obtain controlled monopulse patterns for an adaptive radar. The classical STAP architecture is shown in Figure 1 [2,3]. Our procedure is to first form the adapted sum beam using the classic weight vector [4]
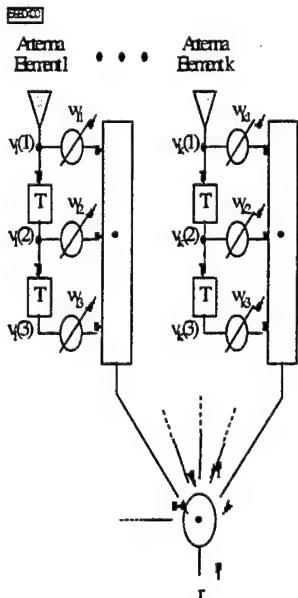
Figure 1. Space-Time Adaptive Processor

$$w = \frac{\Phi^{-1} s^*}{s^T \Phi^{-1} s^*} \qquad (1)$$

where for K antennas and M time taps $w^T = [w_{11} \ldots w_{1M} \ldots w_{N1} \ldots w_{NM}]$, Φ is the NM x NM interference covariance matrix and s is the steering vector for the signal, defined as $s^T = [s_{11} \ldots s_{1M} \ldots s_{NM}]$. For a linear array, and a target at azimuth θ_o and Doppler frequency f_o, the components of s are $s_{nm} = exp[i\, k\, x_n \sin \theta_o - i2\pi mf_o T]$, where x_n is the location of antenna n and k is the wavenumber.

The difference beam $\Delta(\theta, f_o)$ is now formed such that the received interference is minimized subject to the constraints that $\Delta(\theta_o, f_o) = 0$ and the ratio Δ/Σ maintains a constant slope at Doppler frequency f, as specified by

$$\frac{\Delta(\theta_o \pm \Delta\theta, f_o)}{\Sigma(\theta_o \pm \Delta\theta, f_o)} = \pm k_s \Delta\theta \qquad (2)$$

where k_s is a slope constant. If we define a difference – beam weight vector w_Δ, and recognize that the adapted difference pattern in the azimuth-Doppler domain (θ, f) is $w_\Delta^T g = g^T w_\Delta$, where $g(\theta, f)$ is an NM x 1 vector with components $g_{nm} = exp(i\, k\, x_n \sin\theta - i2\pi mfT)$, we see that the above three constraints can be written in matrix notation as

$$H^T w_\Delta = \rho \qquad (3)$$

where *

$$H^T = \begin{bmatrix} g^T(\theta_o + \Delta\theta, f_o) \\ g^T(\theta_o, f_o) \\ g^T(\theta_o - \Delta\theta, f_o) \end{bmatrix}, \qquad (4)$$

---

* In order to ensure that no anomalies occur we actually used $|w^T g|$ in Equation (5).

31

$$\rho = k_s \begin{bmatrix} w^T g(\theta_o + \Delta\theta, f_o) \\ 0 \\ -w^T g(\theta_o - \Delta\theta, f_o) \end{bmatrix} \Delta\theta . \qquad (5)$$

The weight vector w_Δ that minimizes the difference beam interference

$$w_\Delta^H \Phi w_\Delta$$

subject to the constraint in Equation (3) is [5]

$$w_\Delta = \Phi^{-1} H^* \left(H^T \Phi^{-1} H^* \right)^{-1} \rho . \qquad (6)$$

In order to illustrate the results, we consider a 13 element linear array with 14 temporal taps per element, designed to detect low-speed targets in heavy ground clutter (clutter-to-noise ratio = 65 dB per element). In this case, the weight vector in Equation (1) produces a sum beam that has an interference-plus-noise power after adaptation that is close to the noise floor for all target speeds V such that $0.05 < V/V_b < 0.95$, where V_b is the radar blind speed. The weight vector given in Equation (6) produces a difference beam with an adapted, interference-plus-noise power close to the noise floor for all target speeds. The adapted monopulse pattern Δ/Σ is shown in Figure 2 for two different target** speeds. Note that the monopulse slope is nearly linear over the entire 3 dB width of the sum beam, as required. If the constraint in Equation (3) is not applied the adapted monopulse pattern is highly distorted!

SUMMARY

We have developed a procedure to synthesize sum and difference patterns for space-time adaptive arrays in such a way that a specified monopulse response can be achieved. The approach is quite general and has been applied to more scenarios than presented here, including the case of an adaptive array with spatial degrees of freedom only (i.e., M = 1). Additional details are available from the author.

REFERENCES

[1] U. Nickel, "Monopulse estimation with adaptive arrays", *IEE Proc.*, Part F, Vol. 140, pp. 303-308, Oct. 1993.

[2] O. Frost, "An algorithm for linearly constrained adaptive array processing", Proc. IEEE, Vol. 60, pp. 926-935, 1972.

[3] R. Monzingo & T. Miller, Introduction to Adaptive Arrays, New York: Wiley, 1980

[4] L. Brennan & I. Reed, "Theory of Adaptive Radar", IEEE Trans AES-9, pp. 237-252, 1973.

[5] B. VanVeen, "Maximum Variance Beamforming", in Adaptive Radar Detection & Estimation (S. Haykin and A. Steinhardt, editors), Wiley (New York) 1992.

Figure 2. Positive Angle Portion of Adapted Monopulse Pattern for Three Constraints

** The processor produces a different weight vector for each target speed (target Doppler) to which it is tuned.

PASSIVE SURFACE/SUBMERGED DISCRIMINATION IN A SHALLOW WATER WAVEGUIDE: SWELLEX-96 EXPERIMENTAL RESULTS

V. Premus

MIT Lincoln Laboratory
244 Wood Street, Lexington, MA 02420-9185
vpremus@ll.mit.edu

ABSTRACT

Passive acoustic discrimination of submarines from surface clutter in the littoral oceanic waveguide is a problem of paramount importance facing the U.S. Navy anti-submarine warfare (ASW) community. Investigators have suggested that fluctuation statistics of the received pressure field may be used to distinguish surface and submerged passive acoustic sources. Simanin [1] proposed a statistical test based on probabilistic models for amplitude fluctuations of ray paths of surface and submerged origin. Wagstaff [2] has proposed a lofargram normalization technique based on the short-time harmonic mean of FFT magnitude. Recently, Premus [3] has developed an approach based on the short-time-scale scintillation of waveguide normal mode amplitudes. Mode scintillation is induced by target interaction with surface waves and waveguide deformation associated with surface wave action. Mode scintillation has been demonstrated to possess distinct surface/submerged discrimination potential for narrowband acoustic sources in a simulated littoral waveguide [3]. In this paper, data from the ONR-sponsored shallow water experiment, SWellEX-96, is used to validate the proposed classification methodology and illustrate the basic phenomenology.

1. INTRODUCTION

It has been shown that the short time-scale scintillation of normal mode amplitudes is a potential discriminant for passive surface/submerged acoustic source classification in a shallow waveguide [3]. The approach is based on a two part hypothesis which states: 1) source depth modulation and waveguide deformation induced by surface or internal waves produces observable temporal scintillation of normal mode amplitudes, and 2) the probability densities of mode amplitude scintillation for surface and submerged source classes are well separated under modest rms depth fluctuation conditions. In this paper, we report on the application of the mode scintillation classifier concept to narrowband vertical line array data collected during the SwellEX-96 experiment. Results show that this new classifier successfully discriminates between narrowband tones projected from a near-surface source and a deeply submerged source, each subject to moderate (1-2 m rms) depth fluctuation. The two projectors were towed simultaneously at depths of 10 m and 60 m in 200 m of water, at a constant range rate, with source/receiver separation varying from 3-7 km during the course of the experiment. The receiver consisted of a bottom mounted vertical line array (VLA) spanning approximately 80% of the water column. The impact of inaccurate environmental assumptions and limited vertical aperture on processor performance and implementation will be discussed. The results provide encouraging evidence that amplitude modulation may exploited for the purpose of passive surface/submerged classification.

2. THE MODE SCINTILLATION DEPTH CLASSIFIER

In the far field of an acoustic source, the complex pressure field at a depth z may be expressed in terms of a superposition of normal modes, $\Phi_m(z)$, given by,

$$p(z; z_s, r_s) = a \sum_{m=1}^{M} \Phi_m(z_s) \Phi_m(z) \frac{\exp(jk_{rm}r_s)}{\sqrt{k_{rm}r_s}}, \quad (1)$$

where z_s is source depth, r_s is source range, k_{rm} is the horizontal wavenumber corresponding to mode $\Phi_m(z)$, and a is a complex, zero-mean random variable with variance σ_a^2, corresponding to a narrowband stochastic source signal. Thus, the pressure field at the array is equal to a superposition of waveguide mode excitations, each of which is proportional to the amplitude of the mode function at the depth of the source. In vector notation, temporal samples of the noise added pressure field, $\mathbf{p}(t)$, may be written as

$$\mathbf{p}(t) = \mathbf{\Phi H}(t) + \mathbf{n}(t), \quad (2)$$

where $\mathbf{\Phi}$ is the $N \times M$ matrix of normal mode functions, $\mathbf{H}(t)$ is the $M \times 1$ vector of temporally varying modal excitations, $a\Phi_m(z_s(t)) \frac{\exp(jk_{rm}r_s(t))}{\sqrt{k_{rm}r_s(t)}}$, $\mathbf{n}(t)$ is an $N \times 1$ noise vector, and N is the number of sensors. Individual mode amplitudes are resolved at the array by forming the inner-product of the array measurement vector with each individual mode function,

$$\hat{\mathbf{H}}(t) = \mathbf{\Phi}^{\dagger} \mathbf{p}(t), \quad (3)$$

where $\mathbf{\Phi}^{\dagger} = (\mathbf{\Phi}^H \mathbf{\Phi})^{-1} \mathbf{\Phi}^H$ and the superscript H denotes conjugate transpose.

The scintillation index, SI_m, for each mode is computed from the estimated modal excitation time series,

$$SI_m = \frac{Var\{| \hat{H}_m(t) |\}}{E\{| \hat{H}_m(t) |\}}. \quad (4)$$

The interval over which the statistics in (4) are calculated must be long enough to span a few cycles of the surface wave modulation for estimate stability. The interval must also be short enough to insure that fine-scale ocean sound speed variability, as well as source amplitude and phase, are relatively stable over the averaging window.

The mode functions corresponding to a typical downward refracting shallow sound speed profile (SSP) are pictured in Figure 1 for a water depth of 100 m and a frequency of 35 Hz. Mode amplitude fluctuations are computed at each FFT interval. A notional timeseries plot of Mode 3 amplitude fluctuation versus time is depicted for both surface and submerged sources. Mode amplitude scintillation is expected to be greatest near a modal zero-crossing, where the slope of the mode functions is usually a maximum. Mode scintillation is smallest near a modal extremum, where the slope of the mode function is nearly zero. The mode functions for many shallow channels at low frequency are nearly sinusoidal, with all modes sharing a common zero-crossing at the pressure-release surface. Thus, a surface source whose depth is modulated by surface wave interaction will exhibit high scintillation relative to the background for **ALL** modes simultaneously. A source at depth with the same magnitude depth perturbations will exhibit very low scintillation relative to the background for at least one mode, due to its expected proximity in depth to at least one modal extremum.

3. SWELLEX-96 EXPERIMENTAL RESULTS

3.1. Experiment Overview

SwellEX-96 was an ONR sponsored matched field processing experiment conducted off the coast of San Diego near Point Loma during May of 1996 [4]. The experiment site is a shallow water environment with an average depth of approximately 200 meters. A bathymetric map of the experiment site, which identifies the source track and the coordinates of the receiver array, is shown in Figure 2. The data which is subject of this analysis was collected during Event S5 of that experiment. During that event, two acoustic sources were simultaneously towed at a constant speed of 5 knots (2.5 m/s) at depths of 10 meters and 60 meters, respectively, between the points labelled COMEX and FINEX over a 90 minute interval. The approximate range to the sensor during this interval is 4-8 km. The bathymetry is approximately range independent between the source and the sensor. The array, denoted SVLA, is tethered to the bottom. It consists of 32 elements spaced at 5 m intervals. The array is not fully spanning, sampling approximately 80 % of the water column. This condition introduces limitations on the number of modes, in particular the higher order modes, that may be individually resolved at the VLA using the inner-product calculation from the previous section.

The source spectrum of each projector was composed of a comb of narrowband tones in the frequency range 49-235 Hz. The tones were stable in frequency and stable in source level to within +/- 1 dB. Source level stability was especially important in order to focus on the waveguide phenomenology and rule out natural fluctuations in source level at the source as a cause of mode amplitude fluctuations during this experiment. Data analysis focused on the subband from 100-150 Hz, which contained shallow source tones at 109, 127, ans 145 Hz, and deep source tones at 112, 130, and 148 Hz.

3.2. Classifier Performance

Figure 3 shows the source depth record for the deep source during the 90 minute source tow of Event S5. We restrict our attention to the 40 minute segment beginning at approximately 30 minutes and continuing until 70 minutes. During this interval the mean depth of the deep source is about 60 m and its fluctuation level lies in the range 1-2 m rms, consistent with the simulation experiment reported in [3]. Note that a shallow source depth of 10 m is sufficiently shallow to approximate a surface source for the purpose of concept validation in this environment. Source depth time series data for the shallow source was not available. It is assumed to have an rms fluctuation level of similar magnitude of that depicted for the deep source.

The bottom panel of Figure 3 illustrates empirically derived histograms for Mode 3 scintillation index over a 10 minute interval. Observe the pdf separation between the submerged source class (solid) and surface class (dashed-dot), as well as their position relative to that of the background scintillation pdf (dotted) calculated from a quiet non-signal FFT bin. These results are consistent with the simulation experiment results reported in [3].

Classification performance of the mode scintillation algorithm during the 40 minute segment of Event S5, for tones within the frequency range 100-150 Hz, is depicted in Figure 4. In this case, the in-situ measured sound speed profile (SSP) for the environment was used to model the waveguide mode functions and compute the test statistic. The top panel shows the lofargram at an individual sensor for this frequency range. The second and third panels show threshold exceedances for the minimum scintillation index taken across all modes. Different threshold settings have been applied to the same raw scintillation output values to produce the two indicator maps. The thresholds for both the surface and submerged indicator maps have been tuned to permit a small amount of false classification activity to pass through. Hits on the surface indicator map (second panel) at the frequencies 109, 127, and 145 Hz represent correct identifications of the tones originating at the surface source. Notice some incorrect classifications of the three submerged tones at 112, 130, and 148 Hz as surface tones around the 15 minute mark of the second panel. These false classifications are due to source turn-off events that occurred during Event S5 and are not source depth fluctuation induced. The mode scintillation classifier would not be penalized for these incorrect classification decisions as they were not phenomenology induced. In practice, a post-processing step such as an M-of-N integration rule would mitigate against such false classification decisions. Additional shallow source turn off events toward the end of Event S5 are responsible for the false surface classification decisions near the 35 minute mark at frequencies 112 and 130 Hz. Hits on the submerged indicator map (third panel) aligned with frequencies 112, 130, and 148 Hz represent correct identifications of tones originating from the deep submerged source. The performance level reflected by these results is very encouraging. The results suggest that the phenomenology underlying narrowband amplitude fluctuations in this SwellEX-96 data set has been correctly interpreted for the purpose of addressing the surface/submerged discrimination problem.

To illustrate the robustness of the mode scintillation technique to imperfect environmental assumptions in this environment, consider the classifier performance shown in Figure 5. Here, an isovelcity assumption of 1500 m/s was made for the entire span of the water column. The bottom was modeled as a half-space with a constant velocity of 1800 m/s. However, as the mode shapes under this assumption do not deviate significantly from those obtained using the actual SSP, the classifier perfomance is observed to change very little under this relatively crude channel model.

4. CONCLUSION

In this paper, we reported on the validation of the mode scintillation classifer introduced in [3] using data from the SWellEX-96 shallow water experiment. The data analysis reveals that the scintillation of mode amplitudes in reponse to source depth modulation and waveg-

uide deformation associated with surface waves can provide a discriminating feature for surface/submerged passive acoustic classification. Due to its reliance on coarse-scale environmental information (i.e. mode shape) rather than fine-scale detail (i.e.phase), the approach appears to be robust to imperfect channel model assumptions. Simulation studies [3] have also shown the approach to be useful for the effective vertical aperture represented by a horizontal line array (HLA) at endfire, with an array length on the order of 500 m for a 100 m channel at 50Hz. Future work is aimed at integrating the contributory evidence provided by the mode scintillation index into an information-fusion-based architecture for passive acoustic source classification.

5. REFERENCES

[1] A. A. Simanin. Possibility of classifying water and surface rays by amplitude fluctuations of the received signal. *Sov. Phys. Acoust.*, 36(6):622–624, Nov.-Dec. 1990.

[2] R. A. Wagstaff. The AWSUM Filter: A 20-db gain fluctuation-based processor. *IEEE J. Oceanic Eng.*, 22(1):110–118, January 1997.

[3] V. Premus. Modal scintillation index: A physics-based statistic for acoustic source depth discrimination. *J. Acoust. Soc. Am.*, 105(4):1–11, 1999.

[4] N. Booth. Swellex-96 Preliminary Data Report. Technical report, Space and Naval Warfare Systems Center San Diego, Acoustics Branch, Code D881, 1996.
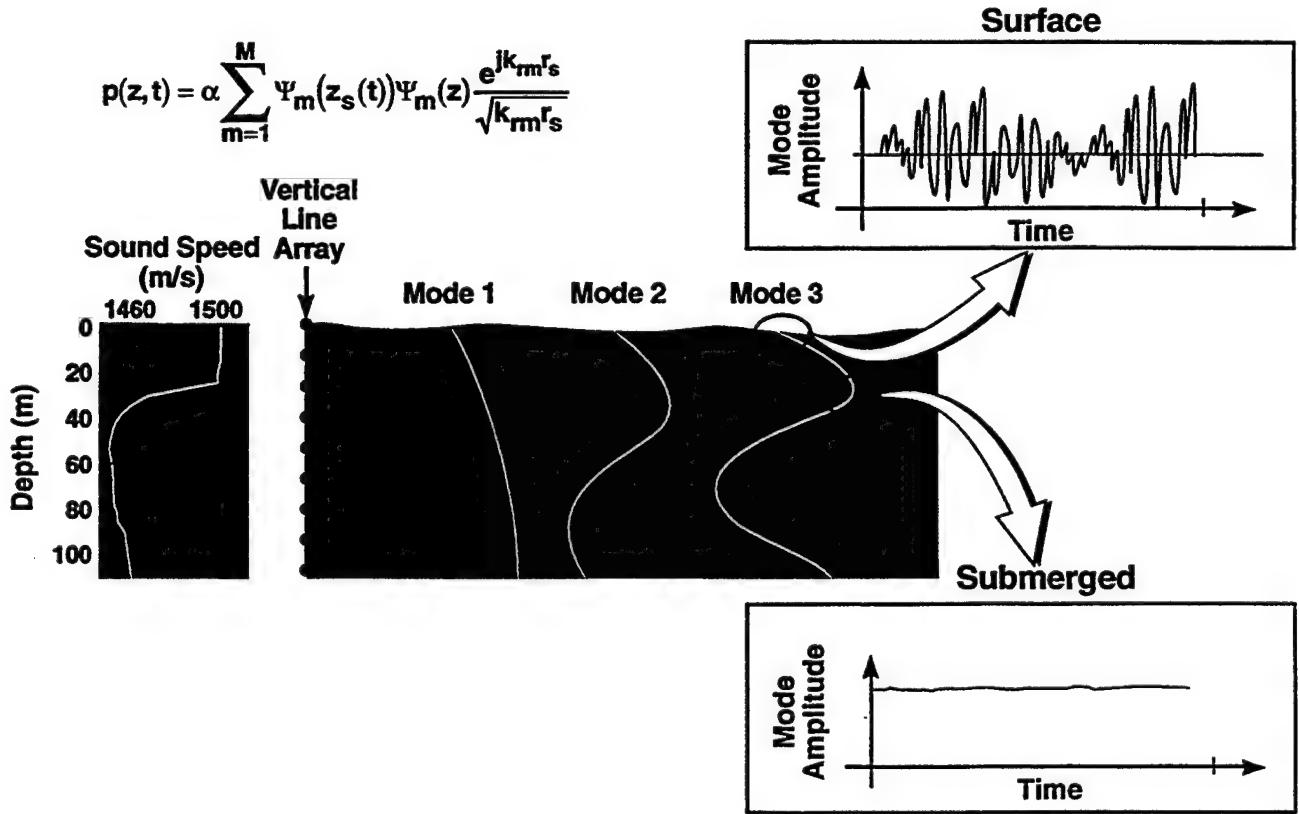
Figure 1: Mode scintillation classifier concept. Mode amplitudes are maximally sensitive to source depth modulation near surface, where the derivative of each mode function is a maximum. Mode amplitudes are minimally sensitive to source depth perturbations near mode maxima..

SwellEx-96 Experiment Overview
Array Geometry
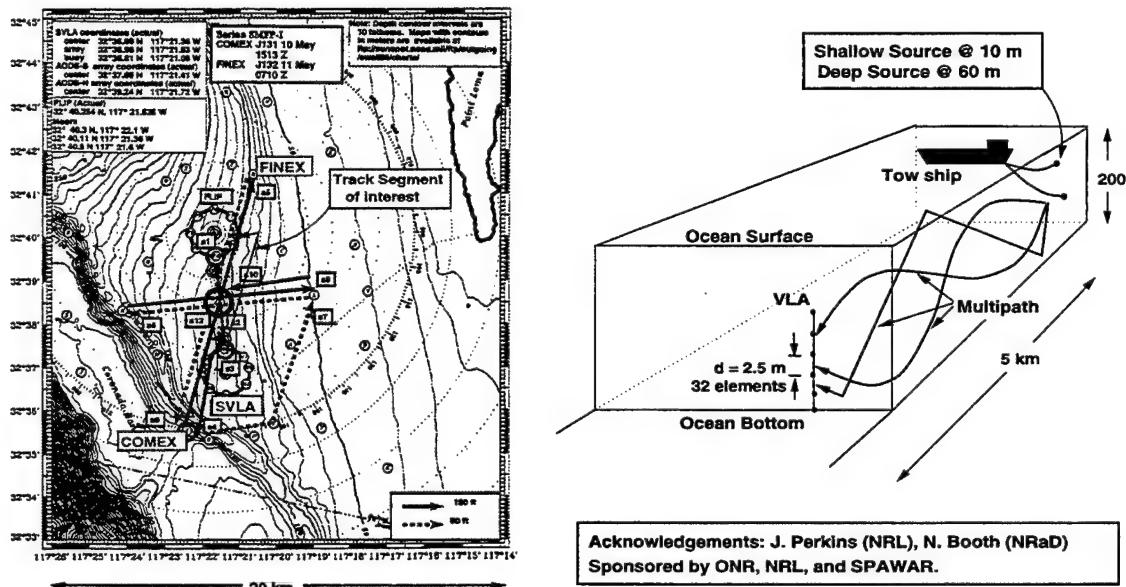
Figure 2: (a) Bathymetric map of SwellEX-96 experiment site. Event S5 source track is highlighted in red. (b) Measurement geometry.

SwellEX-96 Preliminary Results
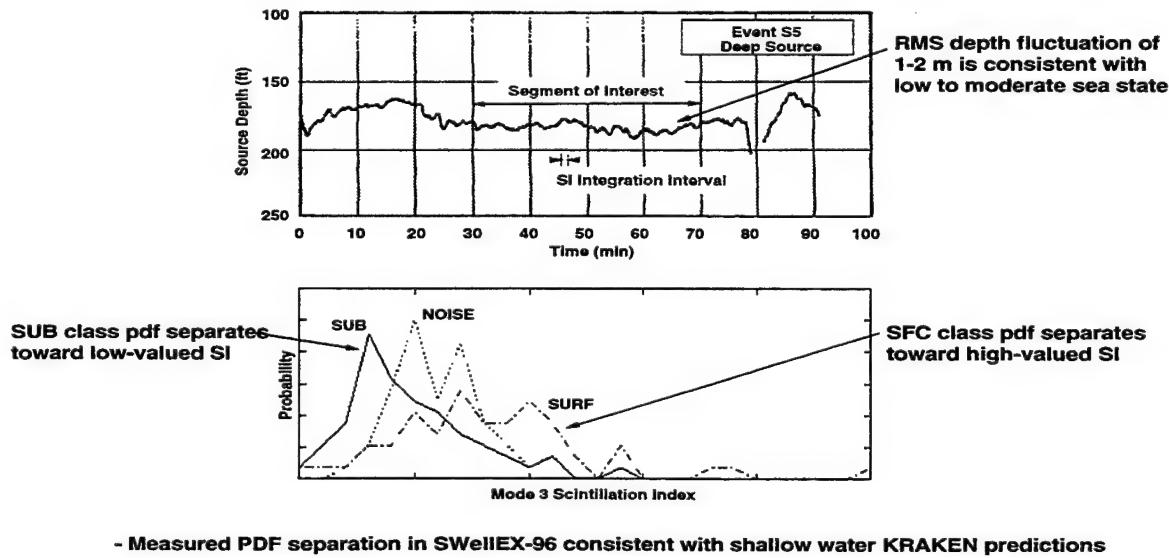
- Measured PDF separation in SWellEX-96 consistent with shallow water KRAKEN predictions

Figure 3: (a) Depth record for deep source during Event S5. The acoustic data used for the classifier analysis was taken from the middle 40 minute segment. During this segment, source depth modulation is approximately zero mean with a standard deviation of 1-2 meters. (b) Empirically derived histograms of Mode 3 scintillation index during 10 minute segment of Event S5.

37

SwellEX-96 Results
Threshholded "Scintillation Grams" Using Measured SSP
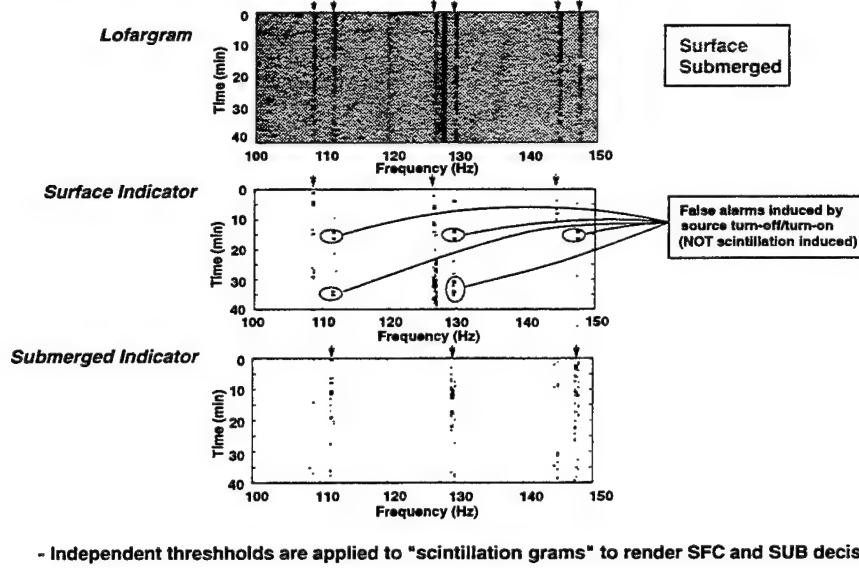
- Independent threshholds are applied to "scintillation grams" to render SFC and SUB decisions

Figure 4: Mode scintillation classifier performance results using actual or measured SSP: (a) Lofargram at a single hydrophone. Surface source tones at 109, 127, and 145 Hz are highlighted with green arrows. Submerged tones at 112, 130, and 148 Hz are highlighted with red arrows. (b) Surface indicator map identifying threshold exceedances consistent with surface source scintillation condition. Note that the false classifications of submerged tones near minutes 15 and 35 are due to deep source turn-off events and are **not** depth modulation induced. (c) Submerged indicator map identifying threshold exceedances consistent with submerged source scintillation condition. Threshold has been tuned to allow a small amount of false classification activity through for visual calibration purposes.

SwellEX-96 Results
Thresholded "Scintillation Grams" Assuming Isovelocity Profile
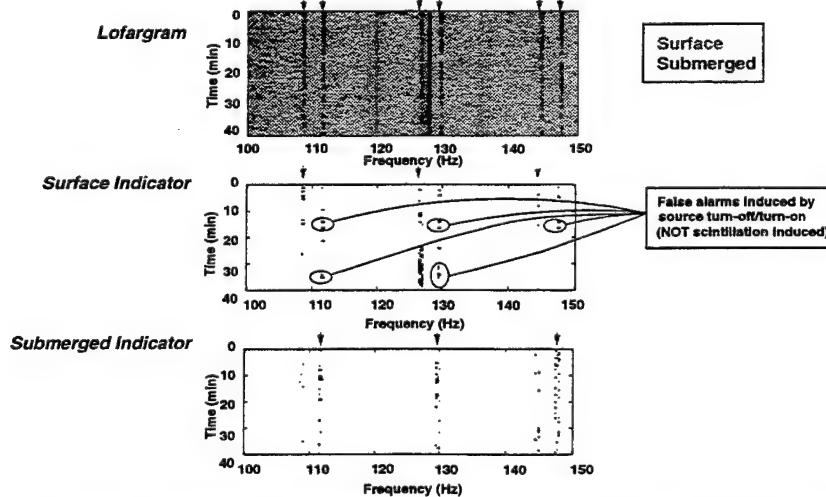
- Emphasis on mode shape rather than fine structure (i.e. phase) enhances tolerance to imperfect environmental model assumptions

Figure 5: Mode scintillation classifier performance results using isovelocity sound speed model. Note the insensitivity of the classifier's performance to this imperfect and relatively crude environmental assumption.

38

TRELLIS STRUCTURE APPROACH TO MULTITARGET TRACKING

Richard Perry   Anand Vaddiraju   Kevin Buckley

Department of Electrical and Computer Engineering
Villanova University
Villanova, PA 19085
E-mail: `perry(anand, buckley)@ece.vill.edu`

ABSTRACT

In this paper we address sequential multitarget tracking for Radar applications. Specifically, we consider the problem of associating location measurement data over time to form estimates of multiple tracks. The algorithmic approach we employ is sequential and very flexible in that it can handle missed detections, false alarms, track initiation, and number-of-track estimation, while employing maximum likelihood or Bayesian cost functions. The approach is based on a trellis diagram which depicts the possible progressions of sequences of location measurements over time. Computation requirements are managed using a generalized K-path, L-list Viterbi algorithm to prune candidate track sets which have prohibitively high costs. Here we focus on a Bayesian solution to the joint problem of estimating both the number-of-tracks and the tracks themselves.

1. INTRODUCTION

Multitarget tracking is an important and challenging problem in radar signal processing [1] which has received intense attention for well over two decades now. Starting from a time evolving set of noisy measurements from detected targets and false alarms, the problem is to associate the detections over time to form multiple target tracks. Some early work focussed on associating time evolving target measurement data, that includes both real detections and false alarms, to one or more target tracks using a likelihood function formulation. Smith and Buechler [2] selected the single track that maximized a likelihood cost – an approach which Morefield [3] extended to maximum likelihood (ML) estimation of multiple tracks. Realizing the need to incorporate prior detection and false alarm probabilities, Singer et al. [4] and Reid [5] developed, respectively, single track and multitrack Bayesian estimators. Both ML and Bayesian multitrack estimators require evaluation of the cost of every possible (candidate) set of tracks through the measurements. The number of candidate track sets increases exponentially with time since at a given time each candidate track can extend to any of the next measurements. To reduce this computation burden, proposed methods typically only evaluate "feasible tracks" in "gating volumes", i.e. only measurements which are in some sense close to a predicted measurement are considered when extending a candidate track to the next measurement time. A result of this approach is that the track which is optimum over the entire processing interval may not be found, and in particular in sequential track estimation valid tracks can be dropped after track-to-measurement association fails for several measurement times.

In [6], an alternative approach to computation reduction is proposed. Therein, a K-path extension of the Viterbi algorithm [7] is used to find the best K nonintersecting tracks according to a deterministic energy cost function. A trellis diagram is defined where at each stage (measurement time) states represents all combinations of measurements taken K at a time. Paths through the trellis represent K track sets. The K-path Viterbi algorithm is used to prune the number of paths through the trellis that must be considered, thereby reducing computation. Globally optimum solutions are computed, thus avoiding the problems associated with feasible tracks and gating volumes. Use of this K-path extension of the Viterbi algorithm is predicated on a "finite state machine structure" or "Markov condition" in the cost function, which is satisfied with the particular deterministic energy cost function used in [6]. However, an ML or Bayesian cost function will not satisfy this requirement since track measurements are correlated throughout the history of a track. (One way to think of this is that, in implementing an ML or Bayesian multitrack estimator, Kalman filters are used in the cost computation to generate measurement innovations. These Kalman filters are recursive and thus have infinite memory.) As a result direct use of K-path Viterbi for multitarget tracking is limited.

For the data communication application, another modification of the Viterbi algorithm, termed list Viterbi [8], has been proposed. This algorithm provides, at each stage through the trellis, a list of best paths through the trellis diagram. In [9], a novel K-track, L list Viterbi algorithm approach was proposed for multitarget tracking. By pruning the trellis to a list of best K-tracks instead of to a single best K-track, this approach overcomes the shortcomings of K-path Viterbi tracking [6] while controlling computation without resorting to the use of gating volumes. It also provides a ranked list of K-track estimates which may, according to the optimality criterion, have very similar costs but represent dissimilar trajectories. Such a list can be valuable for post processing classification and/or data fusion.

The algorithm described in [9]: is based on a Maximum Likelihood (ML) cost function; assumes the number of tracks, K, is known; and requires that the probability of target detection is unity. In this paper we employ a more effective Bayesian cost, describe an algorithm for joint estimation of K and the tracks, and account for missed detections.

Below we first introduce the K target tracking problem. We then develop a Bayesian cost function for estimating K and the tracks, and then marginalize over the tracks to derive an alternative estimator of K alone. The trellis diagram representation of the problem is then constructed, and followed by description of the K-path L-list Viterbi algorithm for Bayesian multiple track esti-

mation. Finally, we illustrate this algorithm with several numerical examples.

2. PROBLEM FORMULATION

In this section we introduce the tracking problem, and we develop notation required to describe the Bayesian estimation approach and the trellis structure upon which the proposed generalized Viterbi algorithm is based.

In the Radar multitarget surveillance problem, the objective is to track multiple target trajectories over time. At each measurement time, measurements are provided from multiple target detections. The measurements for each detection consist of a set of D estimated location and/or velocity parameters. We assume that K, the number of targets, is unknown and constant over the processing time. Detections can correspond to either targets or false alarms. Missed detections will be accounted for. The probability of detection P_d is assumed known, as is the density function for the number of false alarms in the surveillance volume per measurement time. The problem we address here is to estimate K, as $\hat{K}$, while associating the measurements into $\hat{K}$ tracks. The algorithm proposed here, is sequential in that at each measurement time m an estimate $\hat{K}$ and $\hat{K}$ tracks (up to time m) are computed utilizing results from computations of estimates at previous times.

We denote as $z_{m,j}$ the jth the measurement vector at time m. Then, $Z_m = \{z_{m,j}; j = 1, 2, ..., J_m\}$ is the set of J_m measurement vectors at time m, and $\mathcal{Z}^n = \{Z_1, Z_2, ... Z_n\}$ denotes the set of all measurement vectors up to time n.

2.1. Candidate Tracks

A single candidate or postulated track is characterized by a measurement vector or a missed detection for each measurement time. At time n let one such track, the lth track, be denoted θ_n^l. The vector of measurements for this track is

$$\mathcal{Z}(\theta_n^l) = \{z_{1,j_1(l)}, z_{2,j_2(l)}, \cdots, z_{n,j_n(l)}\} \qquad (1)$$

where the subscript $j_m(l)$ is the measurement index at time m for the lth track. We account for the possibility of a missed detection by letting $j_m(l)$ range from 1 to $J_m + 1$, with $j_m(l) = J_m + 1$ indicating a missed detection. Note that a z_{m,J_m+1} is not an actual measurement, but indicates a missed detection at time m, and is a notational convenience. There are $L_n' = \prod_{m=1}^n (J_m + 1)$ of these candidate tracks: $\theta_n^l; l = 1, 2, \cdots, L_n'$.

For a given candidate track θ_n^l, measurement noise is assumed additive, Gaussian and temporally white. The trajectory and measurements are assumed to evolve in time according to the state/measurement equations [1]

$$x_{m+1}^l = \Phi_m x_m^l + w_m^l \qquad (2)$$

$$z_{m,j_m(l)} = H_m x_m^l + u_m^l \qquad (3)$$

where at time m and for candidate track θ_n^l, x_m^l is the state vector, and Φ_m and H_m are the state transition and output matrices respectively. The w_m^l and u_m^l vectors are zero mean, mutually independent, white and Gaussian with known covariance matrices Q_m^l and R_m^l respectively.

---

[1]To simplify the discussion we assume linear state/measurement equations.

A Kalman filter can be applied to this track to smooth the measurements, to provide minimum variance state vector estimates, and to compute the innovations for the measurement sequence $\mathcal{Z}(\theta_n^l)$. The Kalman filter equations are:

Kalman Prediction Equations

$$
\begin{aligned}
\hat{x}_{m+1,m} &= \Phi_m \hat{x}_{m,m} \\
\hat{z}_{m+1,m} &= H_{m+1} \hat{x}_{m+1,m} \\
v_{m+1} &= z_{m+1} - \hat{z}_{m+1,m} \\
\hat{x}_{m+1,m+1} &= \hat{x}_{m+1,m} + G_{m+1} v_{m+1} \qquad (4)
\end{aligned}
$$

Kalman Gain Equations

$$
\begin{aligned}
\Xi_{m+1,m} &= \Phi_{m+1} \Xi_{m,m} \Phi_{m+1}^T + Q_{m+1} \\
S_{m+1} &= H_{m+1} \Xi_{m+1,m} H_{m+1}^T + R_{m+1} \\
G_{m+1} &= \Xi_{m+1,m} H_{m+1}^T S_{m+1}^{-1} \\
\Xi_{m+1,m+1} &= \Xi_{m+1,m} - G_{m+1} H_{m+1} \Xi_{m+1,m} \qquad (5)
\end{aligned}
$$

where G is the Kalman gain matrix, Ξ is the measurement error's covariance matrix, and dependence on the track index l is not shown for notational convenience.

The measurement innovations sequence $\{v_{1,j_1(l)}, v_{2,j_2(l)}, \cdots, v_{n,j_n(l)}\}$ and corresponding covariance matrices $\{S_{1,j_1(l)}, S_{2,j_2(l)}, \cdots, S_{n,j_n(l)}\}$ generated in the Kalman filter computation, are to be used in optimum track estimation. Note that a $v_{m,j_m(l)}$ corresponding to z_{m,J_m+1} is not an actual innovations, but is a result of a missed detection at time m, and again is a notational convenience. In such a case, the Kalman state is just the predicted state (i.e. no measurement is available to update the state).

2.2. Candidate Track Sets

Now consider a set of K tracks. Assume tracks can not share detections, and again assume missed detections are possible. It can be shown that the number of candidate track sets is [2]

$$
I_n' = \left(\sum_{j=0}^K \binom{J_1}{j} \right) \left(\prod_{m=2}^n \left(K! \sum_{j=0}^K \frac{\binom{J_m}{j}}{(K-j)!} \right) \right) \qquad (6)
$$

Let the ith candidate K-track set be represented as

$$
\tau_n^i = \{\theta_n^{l_1(i)}, \theta_n^{l_2(i)}, \cdots, \theta_n^{l_K(i)}\} \quad , \qquad (7)
$$

where the superscript $l_k(i)$ denotes the kth track of the ith track set. $\mathcal{Z}(\tau_n^i)$ will be used to denote the measurement data associated with the ith K-track set.

3. BAYESIAN TRACK ESTIMATION

First we will assume that K is known and consider the problem of estimating K tracks at time n from the measurement data $\mathcal{Z}^n$. We will then extend this to the unknown K problem.

---

[2] $\binom{A}{B}$ represents the combinations of A things, taken B at a time.

3.1. Estimation of K Tracks

Using a Bayesian formulation which incorporates prior track set probabilities, at time n we have the following optimization problem for the selection from $\tau_n^i; i = 1, 2, \cdots I_n'$

$$\max_{\tau_n^i} \quad p\left(\tau_n^i | \mathcal{Z}^n\right) = \frac{p\left(\mathcal{Z}^n | \tau_n^i\right) \cdot p\left(\tau_n^i\right)}{p\left(\mathcal{Z}^n\right)} \quad . \tag{8}$$

$p\left(\mathcal{Z}^n\right)$ is a normalizing factor which serves to make the possible values of $p\left(\tau_n^i | \mathcal{Z}^n\right)$ sum to unity, and can be ignored since it is independent of τ_n^i. $p\left(\tau_n^i\right)$ is the apriori probability of the set of tracks, and is derived from the known probability of missed detections and the false alarm rate. Letting $J_m = T_m^i + F_m^i$, where T_m^i is the number of measurement vectors at time m used in the K track set τ_n^i and F_m^i is the number of assumed false alarms, we have that

$$p\left(\tau_n^i\right) = \prod_{m=1}^{n} P_d^{T_m^i} \left(1 - P_d\right)^{K-T_m^i} p(F_m^i) \, , \tag{9}$$

where $p(F_m^i)$ is the Poisson distributed probability of F_m^i false alarms.

$p\left(\mathcal{Z}^n | \tau_n^i\right)$ is the joint probability density function of the measurements $\mathcal{Z}^n$ conditioned on the track set τ_n^i. For a given τ_n^i, we can partition $\mathcal{Z}^n$ into the data associated with the tracks, $\mathcal{Z}(\tau_n^i)$, and data associated with false alarms, $\overline{\mathcal{Z}}(\tau_n^i)$. We then have

$$p\left(\mathcal{Z}^n | \tau_n^i\right) = p\left(\mathcal{Z}(\tau_n^i) | \tau_n^i\right) \cdot p\left(\overline{\mathcal{Z}}(\tau_n^i) | \tau_n^i\right) \, , \tag{10}$$

where, under the assumption that false alarms are uniformly distributed over the surveillance volume Y,

$$p\left(\overline{\mathcal{Z}}(\tau_n^i) | \tau_n^i\right) = \left(\frac{1}{Y}\right)^{\sum_{m=1}^{n} F_m^i} \quad . \tag{11}$$

In terms of the Kalman filter generated measurement innovations, for the track measurements,

$$p\left(\mathcal{Z}(\tau_n^i) | \tau_n^i\right) = \prod_{p=1}^{K} p(\mathcal{Z}(\theta_n^{l_p(i)})) = \prod_{p=1}^{K} \prod_{m=1}^{n} p(\mathbf{v}_{m,j_m(l_p(i))}) \tag{12}$$

$$= d_n^i \exp\left\{ -\frac{1}{2} \sum_{p=1}^{K} \sum_{m=1}^{n} \mathbf{v}_{m,j_m(l_p(i))}^T \mathbf{S}_{m,j_m(l_p(i))}^{-1} \mathbf{v}_{m,j_m(l_p(i))} \right\} \tag{13}$$

where for each track p the product over m does not include any missed detections since as noted earlier measurements and innovations for missed detections don't exist. Keeping this in mind,

$$d_n^i = \prod_{p=1}^{K} \prod_{m=1}^{n} \frac{1}{(2\pi)^{D/2} \cdot \sqrt{\det(\mathbf{S}_{m,j_m(l_p(i))})}} \quad . \tag{14}$$

The Bayesian cost for estimating the best K tracks is described by (8), along with (9), (11), (13) and (14).

To derive an equivalent cost from which a time recursive trellis structure representation can be derived, take the negative natural log of (8), ignoring $p\left(\mathcal{Z}^n\right)$. For the ith track set, the following equivalent Bayesian optimization problem is obtained:

$$\min_{\tau_n^i} \quad \Lambda^n\left(\tau_n^i\right) = \sum_{m=1}^{n} \Lambda_m(\tau_n^i) \tag{15}$$

where the time incremental cost is

$$\Lambda_m(\tau_n^i) = \lambda_m(\tau_n^i) + \sum_{p=1}^{K} \lambda_m(\theta_n^{l_p(i)}) \tag{16}$$

with

$$\begin{aligned} \lambda_m(\tau_n^i) &= F_m^i \ln(Y) - T_m^i \ln(P_d) \\ &- (K - T_m^i)\ln(1 - P_d) - \ln(p(F_m^i)) \, , \end{aligned} \tag{17}$$

and the pth track incremental cost at time m is

$$\begin{aligned} \lambda_m(\theta_n^{l_p(i)}) &= \frac{1}{2} \ln(\det(\mathbf{S}_{m,j_m(l_p(i))})) + \frac{D}{2} \ln(2\pi) \\ &+ \frac{1}{2} \mathbf{v}_{m,j_m(l_p(i))}^T \mathbf{S}_{m,j_m(l_p(i))}^{-1} \mathbf{v}_{m,j_m(l_p(i))} \, . \end{aligned} \tag{18}$$

Although from (15), the time recursion

$$\Lambda^n(\tau_n^i) = \Lambda^{n-1}(\tau_n^i) + \Lambda_n(\tau_n^i) \tag{19}$$

can be formed, to solve this Bayesian problem directly, at time n I_n' incremental costs must be computed. In the following sections we develop an algorithmic approach which reduces this computational requirement.

3.2. Joint Estimation of K and the Tracks

For the estimation of both the number of tracks K and the tracks themselves τ_n^i, (8) can be considered a function of unknown K, resulting in the following joint estimation problem

$$\max_{K, \tau_n^i} \quad p\left(K, \tau_n^i | \mathcal{Z}^n\right) = \frac{p\left(\mathcal{Z}^n | K, \tau_n^i\right) \cdot p\left(\tau_n^i | K\right) \cdot p(K)}{p\left(\mathcal{Z}^n\right)} \, . \tag{20}$$

If we assume that K is uniformly distributed from $K = 0$ to some maximum value $K = K_{max}$, and that $p\left(\tau_n^i | K\right) = p\left(\tau_n^i\right)$, and that $p\left(\mathcal{Z}^n | K, \tau_n^i\right) = p\left(\mathcal{Z}^n | \tau_n^i\right)$. This results in the estimator

$$\hat{K}, \hat{\tau}_n^i = \arg \min_{K, \tau_n^i} \left\{ \Lambda^n\left(\tau_n^i\right) = \sum_{m=1}^{n} \Lambda_m(\tau_n^i) \right\} \tag{21}$$

where all values are as defined for (15), and now the cost is searched over all $\tau_n^i; i = 1, 2, \cdots I_n'(K)$ for each $K = 0, 1, \cdots, K_{max}$. (Referring to (6), note that the number of candidate track sets for a given K is a function of K.)

3.3. Marginalization for Number of Track Estimation

If estimation of K is of primary interest, it can be estimated without explicit estimation of τ_n^i by marginalizing $p\left(K, \tau_n^i | \mathcal{Z}^n\right)$ over τ_n^i. Compared to (20), the resulting estimator,

$$\hat{K} = \arg \max_K \left\{ p(K | \mathcal{Z}^n) = \sum_{i=1}^{I_n'(K)} p\left(K, \tau_n^i | \mathcal{Z}^n\right) \right\} \, , \tag{22}$$

can have better performance characteristics, which can be better for small number-of-measurement cases. has different performance characteristics, which can be better for small number-of-measurement cases [10].

4. TRELLIS DIAGRAM FORMULATION

Consider the trellis diagram in Figure 4 which depicts the possible progressions of sequences of location measurements over time for $K = 1$. Each stage of the trellis corresponds to a measurement time. For $K = 1$, at the mth stage there are $J_m + 1$ states, one for each measurement and one to account for a missed target. A branch is a connection from a state at stage $m-1$ to another state at stage m. At time n, the L'_n candidate tracks $\theta_n^l; l = 1, 2, \cdots, L'_n$ correspond to the L'_n possible paths through the trellis from stage 1 to n. For each track, we assign an incremental cost to each branch. For example, for the lth candidate track θ_n^l, the branch cost from stage $m - 1$ to m (i.e. from state $j_{m-1}(l)$ to $j_m(l)$) is $\lambda_m(\theta_n^i)$. It is important to note that since for each candidate track a Kalman filter (which has memory back to the 1st stage through the Kalman states) is being used to compute the incremental costs, the cost associated with a branch depends on what track is being considered. That is, each branch has multiple costs assigned to it, one for each track through it. The cost of the path θ_n^l is the sum of its branch costs. The Bayesian track estimation problem is now one of finding the minimum-cost path through this trellis.
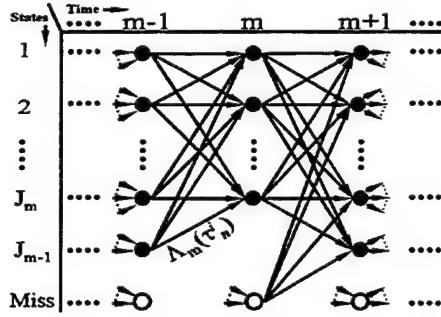

Figure 1: Trellis Diagram for $K = 1$.

For K track estimation, each state of the trellis in Figure 4 represents a set of K measurements and/or missed measurements. So, for stage m with J_m measurements, there are

$$M_m = \sum_{j=0}^{K} \binom{J_m}{j} \tag{23}$$

states. Each branch from stage $m - 1$ to m has a set of $K!/(K - T_m^i)!$ permutations associated with it, specifying the possible orders in which the K measurements and/or missed measurements are used. Now for each candidate K-track set, the incremental cost assigned to a branch is the sum of the costs of K candidate tracks. So, for candidate K-track set τ_n^i, the branch cost from stage $m-1$ to m is $\Lambda_m(\tau_n^i)$. As with the $K = 1$ case, the Bayesian K-track estimation problem is now one of finding the minimum-cost path through this trellis.

5. A LIST VITERBI TRACKING ALGORITHM

Based on the trellis diagram problem formulation described above, we now describe a multitarget track estimation algorithm which provides a list of L best K-tracks as well as estimates of the number of tracks. The algorithm is sequential, self-initiating and can handle false alarms and missed detections. In this section a K-track algorithm is described, so that $K = 1$ is treated within this context.

The Bayesian tracking problem described in Section 3 would, if solved directly, be computationally prohibitive for large n and even moderate values of K and the $J'_m s$. Here we propose to reduce or prune the number of candidate tracks by effectively, at each time m, eliminating from further consideration the candidate tracks that have too large (i.e. too unlikely) total costs at that time. In the trellis problem formulation, this will be implemented by only keeping the L best sets of tracks into each state.

The problem is to find the lowest cost path through the trellis. The Viterbi algorithm [7] can be used to sequentially prune the paths through the trellis for certain ML optimization problems (as well as for some other problems). At any stage in the trellis, it keeps only one path to each state. So it is only applicable to problems for which all other paths can not be optimum. It has been used extensively in digital communication and other application problems for which the sequence can be modeled as a Markov process. It has also been proposed for multitarget tracking based on a deterministic energy cost [6], where the incremental path costs from a state $j_{m-1}(l)$ to a state $j_m(l)$ is a function of only the $j_{m-1}(l)$ and $j_m(l)$ measurements.

In general, multitarget tracking problems, couched in a trellis framework, can not be solved using the Viterbi algorithm. For the Bayesian estimation problem considered here, the need for (infinite memory) Kalman filters for each candidate track precludes its use, so that to solve the problem at time n an exhaustive search of all I'_n paths through the trellis appears necessary. At any stage n in the trellis, this requires that we consider all paths into each state at stage $n-1$, each extended to all states at stage n. That is, the I'_{n-1} paths into stage $n - 1$ must be extended to the I'_n paths into stage n, even though the costs of some of these paths will indicate that the corresponding K-track sets are highly unlikely. Alternatively, we propose to keep, at each stage in the trellis, a "list" of only the best (lowest cost) L paths to each state, where L is selected to assure that no feasible paths are pruned [3].

Generically, the list Viterbi algorithm [8] does this. The multitarget tracking algorithm we propose incorporates the list Viterbi algorithm, along with the K-track trellis formulation of the multitarget tracking problem [6], and Bayesian cost computation based on Kalman filter innovations from Section 3. The algorithm determines a list of L feasible K-tracks sets as a list of L paths through the trellis. In general, paths are evaluated according to a defined cost function, where the incremental costs in progressing from stage to stage are computed as branch costs from states to states.

Algorithm

1) *Setup:* For each time $m = 1, 2, \cdots n$, allocate arrays of size M_m-by-L for the predecessor state indices, predecessor state L-best indices, permutation number of the current state measurement indices to track associations, and current total minimum negative log costs. Allocate Kalman filters for each track in each L-best sets of tracks for each state.

2) *Initialization:* For $m = 1$, $l = 1$, the current costs are set using (17) for the apriori probabilities, since no Kalman fil-

---

[3] Selection of L will depend both on the distribution of false alarm measurements and on the variance of true target measurements. This issue is not addressed here.

ter innovations are available initially. The Kalman filters are initialized using the available measurements for time $m = 1$, and using predicted values for time $m = 2$.

3) *Iteration:* For each time $m = 2, 3 \cdots n$, and for each state $j = 1, 2 \cdots M_m$

    3.1) Add incremental costs (16) from the L-best of each previous state to the total previous costs, and temporarily store these candidate new costs.

    3.2) Select the L-best of the candidate new costs and update the associated Kalman filters.

4) *Results:*

    4.1) Output the final L-best sets of tracks.

    4.2) Output estimates of $\hat{K}$, using just the lowest cost results for a joint estimation (21), or marginalization over the L-best tracks as an approximation to (22).

The algorithm uses a fixed value of K, and can extract k-track sets from the L-best K-track results for $k = 1, 2, \cdots K$.

Although the algorithm is described above in block form, it is time-recursive, and at any time during the iterations can produce current estimates of the L-best tracks and $\hat{K}$. As usual with the Viterbi algorithm, the trellis can be truncated periodically, to limit storage to a fixed size.

6. ILLUSTRATIVE NUMERICAL EXAMPLES

First, two parabolic target tracks consisting of $n = 21$ $\mathcal{X}$ and $\mathcal{Y}$ position values were generated. Gaussian measurement noise with a variance of 0.01 was added to the target tracks. "False detect" events were generated using a Poisson distribution with a false alarm rate of 2, and the false detections themselves were generated using a uniform distribution over the range [0,4]. Probability of a missed detection was assumed to be 0.3, and for the purpose of illustrating that our algorithm can "regain" a target even after a series of missed detections, the lower track was forced to have missed detections from $\mathcal{X} = 1.5$ to 2.5. Other simulation parameters were $K=2$ and $L=16$. Figure 2 shows the $\mathcal{X}$ and $\mathcal{Y}$ position values for the true tracks, and the noisy measurements used as input to the algorithm.
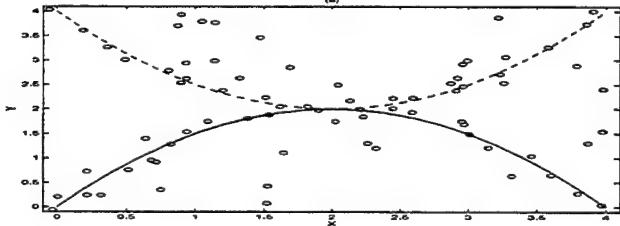
Figure 2: True Tracks and Measurements

Figure 3 shows the best paths, the total cost of which is 117.35. Notice that the one track starts out following the lower target, but switches to the upper target around the region where the true tracks intersect. Figure 4 shows the fourth best paths, which have a total cost of 118.41, just slightly higher than the cost of the best paths, and which correctly follow both the targets. Notice that during missed-detect events, the lower track follows a path tangential to

the true track, but "regains" the target again when detections are recorded.
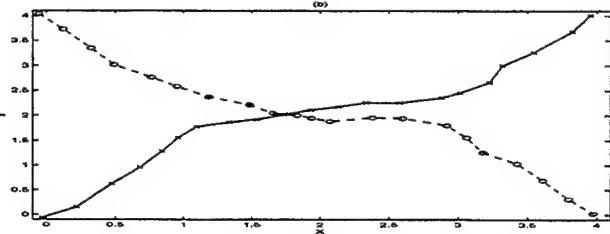
Figure 3: Best set of two tracks, cost=117.35
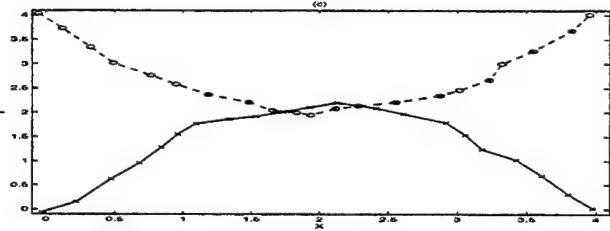
Figure 4: 4th best set of two tracks, cost=118.41.

Second, Monte Carlo simulations were conducted to compare the Bayesian (21) and marginalization (22) estimates of K. N=25 trials were run per measurement noise value. F $K = 2$ linear tracks were simulated, with $n = 6$ measurement times. Measurement noise variance was 0.01. $P_d = 0.7$. Poisson distribution with a false alarm rate of 2, and the false detections themselves were generated using a uniform distribution over the range [0,4]. $K_{max} = 4$ and $L = 16$. Figure 5 shows estimator performance for varying measurement noise variance. On the basis of these simulations, it can be concluded that estimation by marginalization can be advantageous over Bayesian estimation.
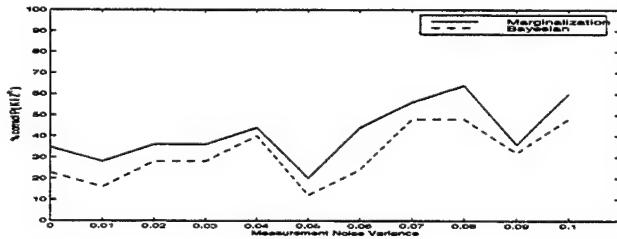
Figure 5: Percentage correct estimates of K.

Last, Monte Carlo simulations were run to study K estimator performance for varying probability of detection P_d. N=100 trials were run per P_d value. $K = 2$ linear tracks were simulated, with $n = 21$ measurement times. Measurement noise variance was 0.01. Poisson distribution with a false alarm rate of 1, and the false detections themselves were generated using a uniform distribution over the range [0,4]. $K_{max} = 3$ and $L = 32$. Table 1 shows the percentages for different estimates of K vs. P_d.

$\rightarrow K$ / $\downarrow P_d$	0	1	2	3
1.00	0	7	93	0
0.95	0	0	96	4
0.90	0	0	91	9
0.85	0	3	85	12
0.80	1	6	72	21

Table 1: Percent correct for the marginalized K estimator.

7. CONCLUSION

In this paper we describe a sequential K-track L-list Viterbi algorithm for Bayeasian measurement /track association and estimation of the number of target tracks K. This algorithm can handle missed detections, false alarms and track initialization. Compared to direct computation of the solution of the Bayesian problem, this algorithm manages computational cost by sequentially eliminating from further consideration candidate K-track sets which have too great a Bayesian cost. The trellis diagram framework onto which the algorithm is built is generally applicable in that: it can accommodate other Bayesian an ML tracking problem solutions, as will as practical approximations to these; and it can be used for situations other than the one addressed here (e.g. for sequential track initiation and elimination).

We plan to further develop this generalized Viterbi tracking algorithm approach, by extending the algorithm to implement sequential track initiation and elimination, by considering further computation reduction via additional candidate K-track set pruning, and by incorporating practical Bayesian/ML motivated techniques such as joint probabilistic data association filtering (JPDAF) [11], probabilistic multi-hypothesis tracking (PMHT) [12] and interactive multiple model (IMM) [13].

8. REFERENCES

[1] Y. Bar-Shalom. Tracking methods in a multitarget environment. *IEEE Trans. on Automatic Control*, Vol. 23, No. 4, pp. 618-626, August 1978.

[2] P. Smith and G. Buechler. A branching algorithm for discriminating and tracking multiple objects. *IEEE Trans. on Automatic Control*, Vol. 20, No. 1, pp. 101-104, February 1975.

[3] C. Morefield. Application of 0-1 integer programming to multitarget tracking problems. *IEEE Trans. on Automatic Control*, Vol 22, No. 3, pp. 302-312, June 1977.

[4] R. A. Singer, R. G. Sea and K. M. Housewright. Derivation and evaluation of improved tracking filters for use in dense multitarget environments. *IEEE Trans. on Info. Theory*, Vol. 20, No. 4, pp. 423-432, July 1974.

[5] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. on Automatic Control*, Vol. AC-24, No. 6, pp. 843-854, Dec. 1979.

[6] A. M. Viterbi J. K. Wolf and G. S. Dixon. Finding the best set of k paths through a trellis with application to multitarget tracking. *IEEE Trans. on Aerospace & Elect. Systems.*, Vol. 25, No. 2, pp. 287-296, March 1989.

[7] G. D. Forney. The viterbi algorithm. *Proc. of the IEEE*, Vol. 61, No. 3, pp. 268-278, March 1973.

[8] N. Seshadri and C. Sundberg. List viterbi decoding algorithms with applications. *IEEE Trans. on Communications*, Vol. 42, No. 2/3/4, pp.313-323, Feb/March/April 1994.

[9] R. Perry, A. Vaddiraju and K. Buckley, "Multitarget list Viterbi target tracking," *Proc. Asilomar Conf. on Circuits, Systems and Computers*, Nov. 1998.

[10] B. Radich and K. Buckley, "Bayesian marginal model selection for low rank sources," *Proc. IEEE Workshop on SSAP*, Sept. 1998.

[11] T. Fortmann, Y. Bar-Shalom and M. Scheffe, "Sonar tracking multiple targets using joint probabilistic data association," *IEEE Jour. of Oceanic Eng.*, Vol. 8, No. 3, pp. 173-183, July 1983.

[12] R. Streit and T. Luginbuhl, "Maximum likelihood method for probabilistic multi-hypothesis tracking," *Proc. SPIE*, Vol. 2335, pp. 394-405, April 1994.

[13] H. Blom and Y. Bar-Shalom, "The interactive multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. on Automatic Control*, Vol. AC-33, No. 8, pp. 780-783, Aug. 1988.

REAL CFAR: NONPARAMETRIC CFAR DETECTION USING REAL TRAINING DATA

E. C. Real

Sanders, A Lockheed Martin Company
P.O. Box 868
Nashua, NH 03061-0868
USA
edward.c.real@lmco.com

D. W. Tufts

Department Of Electrical Engineering
University of Rhode Island
Kingston, RI 02881
USA
tufts@ele.uri.edu

ABSTRACT

We present a method for estimating threshold values for signal detection and classification systems in which a prescribed value of false alarm probability is needed. The threshold values are determined directly from real observed test statistic data without knowledge of the probability distribution of the data.

1. INTRODUCTION

The ability to accurately estimate constant false alarm rate (CFAR) thresholds in an environment in which the distribution of the noise or clutter is both unknown and possibly varying in time, space, or frequency is a critical requirement in many systems. Difficulty in dealing with these issues often leads to ad-hoc methods that are unsatisfying theoretically and can breakdown unexpectedly. To avoid these problems we propose a new method for estimating CFAR thresholds that is based on tolerance interval analysis [1]-[7]. Strict application of the original results of Wilks [2], Wald [3], Scheffé [4] and Tukey [4, 5, 6] requires statistically independent, real valued samples of the data. However, we show that for some applications in signal processing, the requirement for independent samples may be relaxed if the user is willing to accept conservative estimates of the false alarm rate.

The value of tolerance interval analysis was pointed out to us by Roy Streit. A use of tolerance interval analysis in a classification problem has been presented by Streit and Luginbuhl [8].

2. ESTIMATING THE THRESHOLD

Let n represent the number of real independent training samples chosen from the available data. Let X denote this set of n independent samples. Let $X_{(r)}$ be the r-th order statistic of X, where $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$. From [1] we know that for continuous distributions the probability that the quantile ξ lies between the order statistics $X_{(r)}$ and $X_{(v)}$ where $X_{(r)} < X_{(v)}$ is:

$$P(X_{(r)} < \xi < X_{(v)}) = \sum_{k=r}^{v-1} \binom{n}{k} F(\xi)^k (1 - F(\xi))^{n-k} \quad (1)$$

where $F(x)$ is the cumulative distribution function for the unknown distribution.

Similar results can be derived for discrete distributions (e.g. [6, 7]); however in this paper we will limit ourselves to considering only those cases in which data is drawn from continuous distributions.

Using this general result we can compute the specific probability that the quantile ξ lies between any two consecutive order statistics by letting $r = 1, 2, \ldots, n-1$ and $v = r+1$. For example, if we select $n = 100$, $\alpha = 1 - F(\xi) = 0.05$ and assuming we have a continuous distribution we can use equation (1) to obtain the result shown in figure (1).[1]
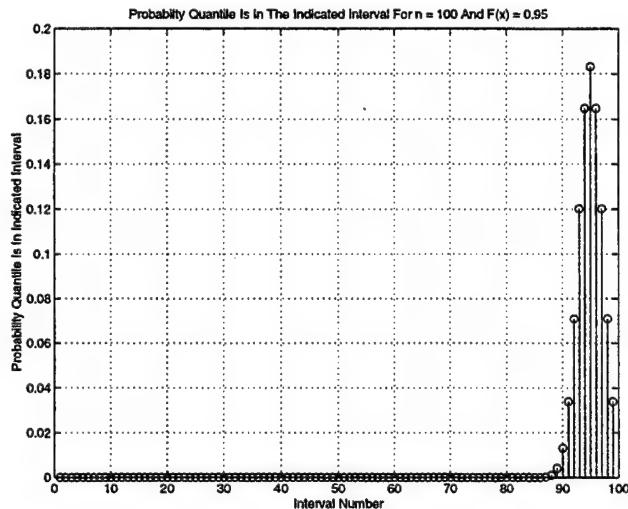
Figure 1: $P(X_{(r)} < \xi < X_{(v)})$ for $r = 1, 2, \ldots, n-1$ and $v = r + 1$ when $n = 100$ and $\alpha = 0.05$.

As indicated, the quantile ξ is most likely in the interval between the order statistics $X_{(95)}$ and $X_{(96)}$ with probability $P(X_{(95)} < \xi < X_{(96)}) = 0.1830$.

For this special case of computing the probabilities that the quantile ξ lies between any two consecutive order statistics, we have a simple formula for determining the interval

---

[1]Note that specifying α is equivalent to specifying $F(\xi)$ so that knowledge of $F(x)$ is unnecessary.

which has the largest probability [9]. This interval is computed as:

$$I = [(n+1)F(\xi)] = [(n+1)(1-\alpha)] \qquad (2)$$

where $[x]$ is the integer portion of x.

If $(n+1)(1-\alpha)$ is an integer then I identifies the second of two consecutive intervals which will have the largest (but equal) probability of containing ξ. That is, if $(n+1)F(\xi)$ is an integer then the intervals specified by $I-1$ and I will have equal probabilities and this probability will be larger than that of any other interval.

The quantile ξ is the threshold we seek. It is the threshold that will achieve the desired false alarm rate. The steps required for estimating this threshold may now be enumerated.

1. We make an initial determination of the number of independent noise or clutter samples required to estimate ξ for a given false alarm rate α as explained in section 3.

2. Given this estimate we select intervals defined by the order indices r and v and compute $P(X_{(r)} < \xi < X_{(v)})$ for every selected interval using equation (1). Note that if we choose the intervals defined by $r = 1, 2, \ldots, n-1$ and $v = r+1$ as in the example above, we may then use equation (2) to determine, in one step, the interval which has the largest probability. If this strategy is used, then the quantity $(n+1)(1-\alpha)$ should be checked to determine if it is an integer and appropriate action should be taken in the following steps if it is.

3. Using either equation (1) or (2), we adopt a maximum likelihood point of view, and select the interval or intervals associated with the largest probability as being the most promising to base our estimate of ξ on.

4. We collect n independent noise or clutter samples, order them, extract those ordered samples which constitute the selected interval and with them compute our estimate of ξ.

To illustrate, in the example given above we found that the maximum interval was bounded by $X_{(95)}$ and $X_{(96)}$. In this case we might estimate ξ by averaging $X_{(95)}$ and $X_{(96)}$ or by simply choosing one or the other of these order statistics as our estimate.

3. ESTIMATING THE SAMPLE SIZE

In this technique, the accuracy of the result is directly related to the number of independent samples used in the computation. We can estimate the number of samples required in the following way. For continuous distributions the probability that all of the independent samples are less than the quantile ξ is:

$$P(X_{(n)} < \xi) = F(\xi)^n \qquad (3)$$

Let $\alpha = 1 - F(\xi)$ be the desired false alarm rate. To insure the proper functioning of our algorithm, we want our chosen quantile to lie within the range of our independent samples. That is we wish $X_{(1)} < \xi < X_{(n)}$. Therefore we choose n so that the probability given by equation (3) poses an acceptable level of risk.

For example, if the desired false alarm rate α is five percent, then for $n = 100$ independent real samples drawn from a continuous distribution the probability that all of the samples are less than ξ is $F(\xi)^n = 0.95^{100} = 0.0059$. If this probability constitutes an acceptable level of risk then we may be satisfied with 100 independent samples. If this is not the case then we can increase n until the risk is acceptable.

More directly, from equation (3) above, let $R = P(X_{(n)} < \xi) = F(\xi)^n$ be a specified level of risk that the true false alarm threshold is greater than the largest order statistic $X_{(n)}$. We then estimate the required number of samples as:

$$n = \lceil \log(R)/\log(1-\alpha) \rceil \qquad (4)$$

where $\lceil x \rceil$ indicates that x is to be rounded up to the next integer.

It is assumed above and throughout this paper that any estimate of n is also consistent with estimates of the time or space intervals over which the data can be assumed to be stationary. If these stationary intervals (in samples) are longer than n then the algorithms presented here will work well. If this is not the case, then we may decide to reduce the sample size n to the point that it becomes consistent with these intervals. If following this, the risk R becomes too high, then we may decide to suffer the effects of mixing data drawn from statistically different distributions which result from using a larger value of n, or we may choose another approach to estimating the false alarm rate threshold. However, we note that any other approach based on sampled data must also address these same trade-off issues.

4. ESTIMATING A CFAR THRESHOLD

The method of estimating the false alarm rate threshold ξ outlined above assumes a stationary distribution. However, in many real world cases the distribution of the environmental noise or clutter is not only unknown, but non-stationary as well. Fortunately, the above technique is flexible enough to accommodate these cases. The steps in the algorithm are as follows.

1. We determine the number of independent samples and particular order statistics required to determine the desired threshold as outlined in sections 2 and 3.

2. We estimate a sample interval (in time or space, depending on the application) such that samples separated by more than this amount can reasonably be assumed to be independent. For example, in typical time domain applications the CFAR threshold is applied after some filtering operation. In these cases one could estimate the equivalent noise bandwidth of the filter and down sample the filter output by the inverse of this bandwidth in samples.

46

3. At predetermined intervals of time or space n independent samples are collected, sorted, the order statistics are chosen, and a new threshold ξ is computed and applied to the input data stream.

An example of a sliding (lag) window version of this algorithm is show in figure (2).

Note that once a target signal is detected, the processing described above must be altered to account for the signal energy that is being added to the data stream. One approach is to simply hold the current threshold until the signal has passed. This approach assumes that the signal energy is reasonably constant over the duration of time the signal is present. However if the signal is present for an extended period of time, and its energy is also changing over that time, we would like to be adaptive to that situation. In this case, one strategy is to construct a tolerance interval constant false *rejection* rate algorithm in a manner similar to the one described here and use it to determine when the signal has passed. In this case the threshold estimate would (most likely) be computed from the *smaller* order statistics and individual samples from the data stream would be tested to see if they fell *below* the estimated threshold.

Other applications, such as automatic gain control (AGC) problems, and other variations on the basic algorithm are easily imagined.
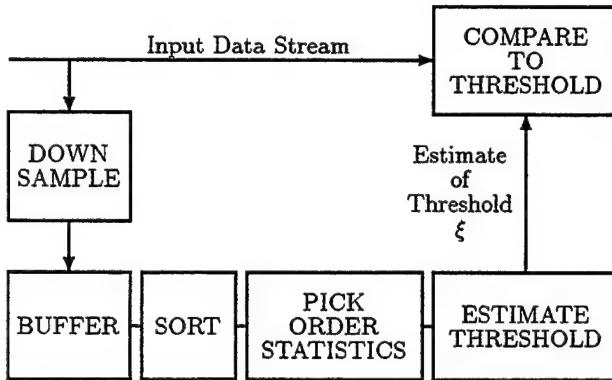

Fig 2. A Tolerance Interval CFAR Algorithm

5. AN EXAMPLE USING INDEPENDENT SAMPLES

In order to illustrate the performance of CFAR algorithms designed in this manner we select the algorithm shown in figure (2) as an example. The performance of this algorithm is demonstrated first in figure (3) by superimposing the computed time varying CFAR thresholds on top of the input data stream. For this example the input data are independent samples of simulated non-Gaussian clutter. The independence of the input samples eliminated the need for down sampling so this was not done. The design probability of false alarm was chosen to be five percent. A buffer length of 100 samples was used, and with no down sampling this became a sample by sample sliding window of that length. The order statistic intervals required in step (2) of the algorithm design sequence given in section 2 were chosen to

be $r = 1, 2, \ldots, 99$ and $v = r + 1$.

These selections result in the same quantile probability graph as that shown in figure (1). As noted earlier we see that the threshold ξ is most likely in the interval between order statistics $X_{(95)}$ and $X_{(96)}$ with probability $P(X_{(95)} < \xi < X_{(96)}) = 0.1830$ (continuous distribution assumption). For simplicity we choose the estimate of the threshold ξ to be the 96-th order statistic $X_{(96)}$. On average, a selection such as this should produce a slightly conservative estimate of ξ. That is one that produces a false alarm rate slightly less than the designed false alarm rate of five percent. The observed false alarm rate of the 10,000 samples of the clutter shown in figure (3) was 0.0508 (5.08 percent) in good agreement with the target design.
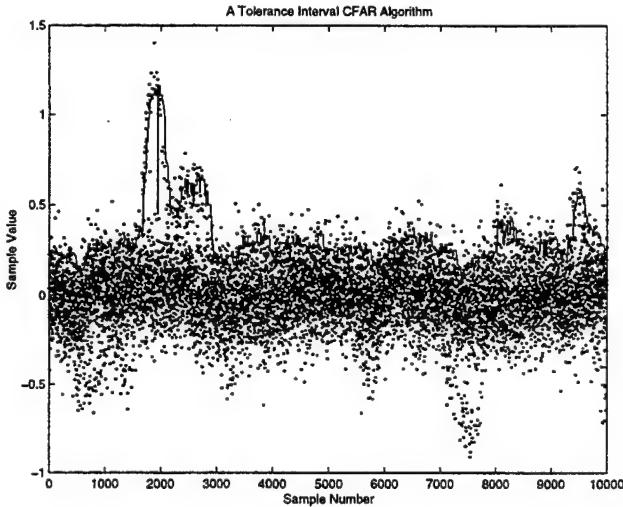
Figure 3: Typical performance of the tolerance interval CFAR algorithm of figure (2) in simulated clutter. The threshold value as a function of sample number is represented by a solid line while the data values are represented by dots.

6. AN EXAMPLE USING CORRELATED SAMPLES

In this section we attempt to demonstrate the utility of our approach when it is applied to non-independent samples. In figure (4) we see a magnitude plot of a sonar return. No target is present in the return. The data is non-Gaussian, non-stationary and correlated. In order to illustrate the effect of using non-independent samples, no attempt is made to down sample or otherwise decorrelate the data.

A further modification to the algorithm shown in figure (2), is the introduction of a split window data buffer. In this version of the algorithm, the data making up the buffer consists of samples that are temporally on either side of the sample to be tested. Half the samples making up the buffer come from a time before the test sample and half come from after. In practice, this introduces a sample processing delay of half the buffer length in the computation and application of the CFAR threshold. The benefit of this approach is that a split window buffer takes better advantage of the statistics

47

of the data surrounding the test sample than a simple lag window does; and thus generally produces more accurate results.

The resulting adaptive threshold can be seen in figure (4) riding along the top of the data. Note that because of the split window implementation, the threshold rises in anticipation of upswings in the data. For this example, the designed probability of false alarm α was 0.01, and the risk factor R was set to 0.01. From equation (4) we calculate the number of independent samples that are required for the proper functioning of the algorithm to be 459. To make the sub-buffers before and after the test sample the same length we round this number to 460, making each sub-buffer 230 samples long. With n set to 460 we use equation (2) to compute the indices of order statistics bounding the maximum likelihood interval for the desired quantile. These are calculated to be the 456-th and 457-th order statistics. These order statistics are averaged to produce the estimate of the quantile. The example shown in figure (4) consists of 12,000 samples. The measured probability of false alarm for this example was 0.0056.

This lower than expected false alarm rate is characteristic of the algorithm's performance when it is applied to correlated data. This result is a consequence of the assumption, made in the derivation of the algorithm, that the samples are all independent. Typically the effect is slight, but the more highly correlated the data, the greater the error. For example, in the extreme case where all of the measurements are the same value, application of the algorithm will result in a probability of false alarm of zero, no matter what probability of false alarm is chosen in the design.
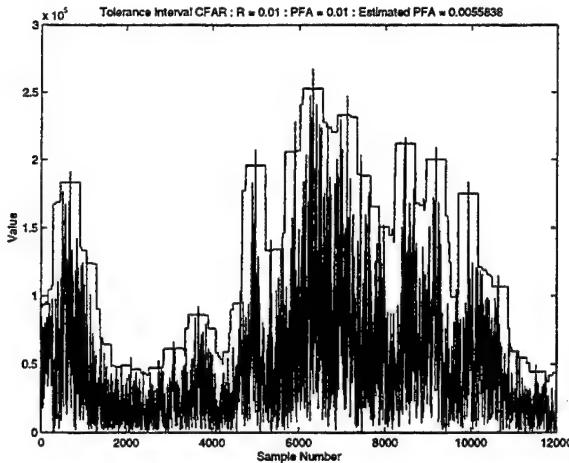
Figure 4: Sonar data with a split window threshold applied. Designed PFA = 0.01. Measured PFA = 0.0056.

7. LOOSING DETECTIONS

Although the algorithm for estimating an adaptive CFAR threshold that we present here has good accuracy with respect to the designed for PFA, there is a danger in overemphasizing the achievement of low false alarm rates in a single stage of processing. The danger is that detections may be lost, especially when the data is highly non stationary. An example is shown in figure (5). In this figure we once again show a magnitude plot of non-stationary sonar returns with the split window adaptive CFAR threshold estimate floating on the top of the data. Here the designed for PFA α was 0.001, the risk factor R was set to 0.01, and the number of samples points used in adaptively estimating the threshold was 4604. These sample points were gathered from two split window buffers of length 2302 that straddled the test sample. The order statistics bounding the maximum likelihood tolerance interval were computed via equation (2) to be the 4600-th and 4601-th. These order statistics were averaged to produce the threshold estimate. This resulted in an estimated PFA of 0.00033 based on 12,000 samples.

We see from this experiment that, although conservative due to the correlated data, the estimated PFA is in good agreement with the designed for PFA of 0.001. However in examining figure (5) we see that there are several regions where there is a significant probability of missed detections. This is due to the non-stationarity of the data. In using a sample support of 4604 samples we are exceeding the interval over which the data is stationary, and "blending" different statistical regions. This results in the algorithms "seeing" the rise in the amplitude (due to the split window) too early, and consequently it prematurely raises the threshold in response. Similarly, the algorithm "remembers" the high amplitude for too long, maintaining the high threshold. A lag window would not raise the threshold prematurely, but would "remember" the high amplitude for twice as long.

This example illustrates that the algorithm presented here has the capability of achieving any desired false alarm rate design; but emphasizes the importance of not over controlling false alarm rates too early in the processing chain, lest detections are missed.
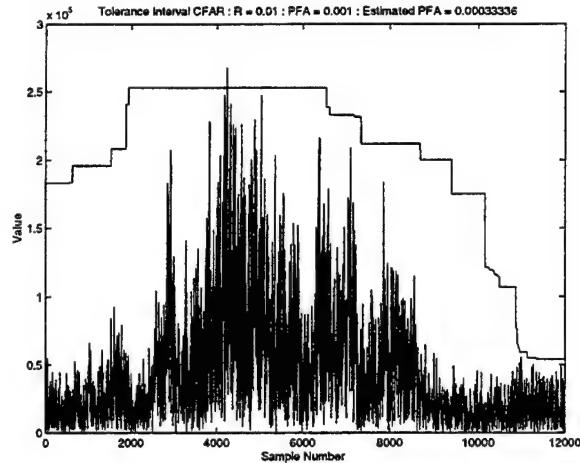
Figure 5: Sonar data with a split window threshold applied. Designed PFA = 0.001. Measured PFA = 0.00033

8. CONCLUSIONS

We have presented a simple, general and flexible method for computing adaptive and non-adaptive CFAR thresholds

based on tolerance interval analysis. Implementation requires few or no arithmetic operations. The use of tolerance intervals eliminates the requirement that the distribution of the input data must be known. Further, the only requirements placed on the data are that the data be real valued and that the samples upon which the analysis is based be independent. Specific examples of the construction of an adaptive CFAR algorithm were given and its performance was demonstrated.

9. ACKNOWLEDGEMENTS

10. REFERENCES

[1] John W. Pratt and Jean D. Gibbons, *Concepts of Non-parametric Theory*, Springer-Verlag, 1981.

[2] S. S. Wilks, "Determination of Sample Sizes for Setting Tolerance Limits", *Annals of Mathematical Statistics*, Vol. 12, pp. 91-96, 1941.

[3] Abraham Wald, "An Extension of Wilks' Method for Setting Tolerance Limits", *Annals of Mathematical Statistics*, Vol. 14, pp. 45-55, 1943.

[4] H. Scheffé and J. W. Tukey, "Non-Parametric Estimation. I. Validation of Order Statistics", *Annals of Mathematical Statistics*, Vol. 16, pp. 187-192, 1945.

[5] John W. Tukey, "Non-Parametric Estimation II. Statistically Equivalent Blocks and Tolerance Regions - The Continuous Case", *Annals of Mathematical Statistics*, Vol. 18, pp. 529-539, 1947.

[6] John W. Tukey, "Non-Parametric Estimation, III. Statistically Equivalent Blocks and Multivariate Tolerance Regions - The Discontinuous Case", *Annals of Mathematical Statistics*, Vol. 19, pp. 30-39, 1948.

[7] D. L. Hanson and D. B. Owen, "Distribution-Free Tolerance Limits : Elimination of the Requirement that Cumulative Distribution Functions be Continuous", *Technometrics*, Vol. 5, No. 4, pp. 518-522, November 1963.

[8] Roy L. Streit and Tod E. Luginbuhl, "Maximum Likelihood Training of Probabilistic Neural Networks", *IEEE Transactions on Neural Networks*, Vol. 5, No. 5, pp. 764-783, September 1994.

[9] Athanasios Papoulis, *Probability, Random Variables, and Stochastic Processes - Second Edition*, McGraw-Hill, 1984.

QUADRATICALLY CONSTRAINED RLS FILTERING FOR ADAPTIVE BEAMFORMING AND DS-CDMA MULTI-USER DETECTION

Zhi Tian, Kristine L. Bell, and Harry L. Van Trees

Center of Excellence in C3I, George Mason University
Fairfax, VA 22030-4444, USA

ABSTRACT

The Minimum Output Energy (MOE) detector used in Direct Sequence Code Division Multiple Access (DS-CDMA) communications systems is analogous to the Linearly Constrained Minimum Power (LCMP) beamformer used in array processing. In both applications, a quadratic constraint on the weight vector norm can improve robustness to mismatch in the temporal or spatial signature vector. We describe a technique for implementing a quadratic inequality constraint with Recursive Least Squares (RLS) updating using the Generalized Sidelobe Canceller (GSC) (also called the Partitioned Linear Interference Canceller (PLIC)) structure. A variable loading term is incorporated at each step, where the amount of loading has a closed form solution. Comparisons are made with several other robust RLS and Least Mean Squares (LMS) algorithms. Simulations show that the variable loading RLS technique offers better convergence and robust control over mismatch than other LMS and RLS implementations for both beamforming and multi-user detection.

1. INTRODUCTION

The Minimum Output Energy (MOE) detector used in Direct-Sequence Code Division Multiple Access (DS-CDMA) communications systems [1] is analogous to the Linearly Constrained Minimum Power (LCMP) beamformer used in array processing [2], [3]. The optimal detector/beamformer minimizes the output power subject to one or more linear constraints. It can be implemented in a direct (constrained) form or using the Generalized Sidelobe Canceller (GSC) structure [4] (also referred to as the Partitioned Linear Interface Canceller (PLIC) [5], [6]). Both Least Mean Squares (LMS) [1], [5], [6] and Recursive Least Squares (RLS) [7] adaptive implementations have been proposed. In both the beamforming and communications applications, a quadratic constraint on the weight vector norm can improve robustness to mismatch in the temporal or spatial signature vector

[1], [6], [8]. The detector/beamformer weights that minimize the output power subject to a set of linear constraints and an inequality constraint on the norm of the weight vector have the same form as the MOE/LCMP weights with diagonal loading of the covariance matrix [1], [8]. The amount of diagonal loading is adjusted to satisfy the quadratic inequality constraint, however the optimal loading level cannot be directly expressed as a function of the constraint and has to be solved for numerically. Even if the optimal loading level is known, implementation in a sequential updating algorithm such as LMS and RLS can be difficult.

Robust LMS implementations that have been proposed include direct form with fixed loading [1], [8], direct form with scaled projection (SP) [8], and GSC/PLIC with scaled projection [6]. In fixed loading, a pre-specified loading term is added at each step. The optimal loading level depends on the scenario, thus a fixed loading may be too high or low for a given situation. In scaled projection, the quadratic constraint is imposed at each update by scaling back the norm of the adaptive portion of the weight vector. It is a simple and generally very effective and robust technique. However, it is subject to the limitations of the class of LMS algorithms in terms of convergence and accuracy.

There do not appear to be any similar technique for RLS updating. In the GSC/PLIC structure, an RLS implementation using scaled projection can be derived, however it is not as effective when used with RLS [9]. In [9], we developed a variable loading (VL) approach for implementing a quadratic constraint in RLS-based adaptive beamformers using the GSC/PLIC structure. A variable diagonal loading term is incorporated at each step, where the amount of loading has a closed form solution that is close to optimal. In [9], significant performance improvements over adaptive beamformers using RLS or RLS with scaled projection were obtained under mismatch situations. In this paper, we apply the technique to the DS-CDMA multi-user detection problem. An equivalent RLS implementation using variable loading is developed for a PLIC detector. Comparisons are made with the other robust LMS and RLS algorithms. Simulations show that the variable loading RLS technique offers better convergence and robust control over mismatch

than other LMS and RLS implementations for both beamforming and multi-user detection.

2. DS-CDMA SIGNAL MODEL

Consider a simple M-user, synchronous, DS-CDMA binary communications system operating in an additive white Gaussian noise (AWGN) channel. The received baseband signal can be modeled as

$$x(t) = \sum_{k=-\infty}^{\infty} \sum_{m=1}^{M} A_m b_m(k) s_m(t - kT_s) + \sigma_n n(t) \quad (1)$$

where T_s is the symbol interval, $b_m(k)$ is the kth data bit of the mth user, $s_m(t)$ is the spreading waveform of mth user with support on $[0, T_s]$, A_m is the amplitude of mth user signal, $n(t)$ is AWGN with unit power spectral density, and σ_n^2 is the noise power spectral density. The spreading waveform is defined as

$$s_m(t) = \sum_{l=0}^{L-1} c_m(l) \psi(t - lT_c) \quad (2)$$

where L is the number of chips per symbol (processing gain), $T_c = T_s/L$ is the chip interval, $c_m(l)$ is the lth chip in the spreading sequence of the mth user, and $\psi(t)$ is the chip waveform with support on $[0, T_c]$. It is assumed that the data symbols are independent, equally likely ± 1 random variables which are independent from $n(t)$. The chips in the spreading sequence $c_m(l)$ are ± 1, and the chip waveform $\psi(t)$ is normalized to have unit energy.

The received signal is passed through the chip waveform matched-filter and sampled at the chip rate, yielding an $L \times 1$ vector of samples,

$$\mathbf{x}(k) = \sum_{m=1}^{M} A_m b_m(k) \mathbf{c}_m + \sigma_n \mathbf{n}(k) \quad (3)$$

where $\mathbf{c}_m$ is the mth user's temporal "signature" vector and $\mathbf{n}(k)$ is a zero mean Gaussian random vector with covariance matrix $\mathbf{I}$. Under ideal conditions the signature vector $\mathbf{c}_m$ is the mth user's code vector.

3. OPTIMUM LINEAR DETECTORS/BEAMFORMERS

In the DS-CDMA communications system, a linear detector weights and sums the samples to produce the statistic

$$y(k) = \mathbf{w}' \mathbf{x}(k). \quad (4)$$

A bit decision is made according to

$$\hat{b}_m(k) = \text{sgn}\left(real(y(k))\right). \quad (5)$$

In multi-user environments, different criteria can be used to derive optimum detectors/beamformers. In [1], the minimum output energy (MOE) detector weights were found by minimizing the energy in $y(n)$ subject to the constraint that the inner product of the weight vector with the desired user's code be a constant value, chosen to be one, i.e.

$$\min \mathbf{w}' \mathbf{R_x} \mathbf{w} \qquad st. \mathbf{c}_1' \mathbf{w} = 1. \quad (6)$$

This formulation is identical to the minimum power distortionless response (MPDR) criterion in adaptive beamforming [2], [3]. The optimal weights are

$$\mathbf{w} = \frac{\mathbf{R_x^{-1}} \mathbf{c}_1}{\mathbf{c}_1' \mathbf{R_x^{-1}} \mathbf{c}_1}. \quad (7)$$

The linearly constrained minimum power (LCMP) criterion [2], [3] provides generalization of this result [5]. The weights minimize the output power subject to a set of q linear constraints of the form $\mathbf{C}' \mathbf{w} = \mathbf{f}$, where $\mathbf{C}$ is the $L \times q$ constraint matrix, and $\mathbf{f}$ is the $q \times 1$ vector of constraint values. The optimal weights are

$$\mathbf{w} = \mathbf{R_x^{-1}} \mathbf{C} \left[\mathbf{C}' \mathbf{R_x^{-1}} \mathbf{C} \right]^{-1} \mathbf{f}. \quad (8)$$

An equivalent representation for the constrained detector is the partitioned linear interference canceler (PLIC) [5], also known as the generalized sidelobe canceler [4] in the array processing literature. In this structure, the optimal solution has the form

$$\mathbf{w} = \mathbf{w}_c - \mathbf{B} \mathbf{w}_a. \quad (9)$$

where the vector $\mathbf{w}_c$ is the fixed $L \times q$ constraint weight vector

$$\mathbf{w}_c = \mathbf{C}(\mathbf{C}'\mathbf{C})^{-1} \mathbf{f}, \quad (10)$$

$\mathbf{B}$ is the $L \times (L-q)$ blocking matrix, and $\mathbf{w}_a$ is the $(L-q) \times 1$ adaptive weight vector. In CDMA detection, $\mathbf{w}_c$ is the optimum single-user detector when $\mathbf{C} = \mathbf{c}_1$ and $\mathbf{f} = 1$. When $\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_M \end{bmatrix}$ and $\mathbf{f} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$, $\mathbf{w}_c$ is the zero-forcing or decorrelating detector. This is the equivalent of forcing a set of null constraints on the other users [5].

The adaptive weight vector has the form

$$\mathbf{w}_a = (\mathbf{B}' \mathbf{R_x} \mathbf{B})^{-1} \mathbf{B}' \mathbf{R_x} \mathbf{w}_c. \quad (11)$$

The matrix $\mathbf{B}$ is an $L \times (L - q)$ unitary blocking matrix which is orthogonal to $\mathbf{C}$, i.e. $\mathbf{B}' \mathbf{C} = \mathbf{0}$, and $\mathbf{B}' \mathbf{B} = \mathbf{I}$.

Let $y_c(k) = \mathbf{w}_c' \mathbf{x}(k)$ denote the output of the constrained path and $\mathbf{z}(k) = \mathbf{B}' \mathbf{x}(k)$ denote the output of the blocking matrix. The adaptive weights can also be expressed as

$$\mathbf{w}_a = \mathbf{R_z^{-1}} \mathbf{p_z} \quad (12)$$

where $\mathbf{R_z} = \mathbf{B}' \mathbf{R_x} \mathbf{B}$ is the $(L - q) \times (L - q)$ covariance matrix of $\mathbf{z}(k)$ and $\mathbf{p_z} = \mathbf{B}' \mathbf{R_x} \mathbf{w}_c$ is the $(L - q) \times 1$ cross-correlation vector of $\mathbf{z}(k)$ and $y_c(k)$.

4. QUADRATIC CONSTRAINTS

Robustness to mismatch in the temporal or spatial signature vector can be achieved by limiting the norm of the weight vector [1],[8]. This requires incorporating a quadratic inequality constraint on $\mathbf{w}$ or $\mathbf{w}_a$ of the form $\mathbf{w}'\mathbf{w} \leq T_o$ or $\mathbf{w}_a'\mathbf{w}_a \leq \beta^2 = T_o - \mathbf{w}_c'\mathbf{w}_c$.

In the PLIC/GSC structure, the optimal solution for the quadratically constrained adaptive weight vector has the same form as (12) with a diagonal loading term $\lambda\mathbf{I}$ added to $\mathbf{R}_z$,

$$\mathbf{w}_a = (\mathbf{R}_z + \lambda\mathbf{I})^{-1}\mathbf{p}_z. \tag{13}$$

The weight vector norm is decreasing in λ for $\lambda > 0$. When $\lambda = 0$, the standard adaptive weight vector (Eq. 11) is obtained. When $\lambda \to \infty$, $\mathbf{w}_a = 0$ and the non-adaptive detector is obtained. The amount of diagonal loading is adjusted to satisfy the quadratic constraint, however the optimal loading level cannot be directly expressed as a function of the constraint and has to be solved for numerically. Alternatively, a fixed level of diagonal loading can be incorporated to improve robustness without satisfying the quadratic constraint exactly. The required level of diagonal loading depends on the scenario, therefore a fixed level may work well in some situations but may be too high or low in others. On the other hand, imposing a fixed quadratic constraint seems to work well over a wide range of scenarios. A reasonable constraint value is $T_o = 2\mathbf{w}_c'\mathbf{w}_c$.

5. ADAPTIVE ALGORITHMS

The linearly constrained detector/beamformers can be implemented adaptively using sequential update algorithms such as LMS or RLS in both the constrained and PLIC/GSC structures [1], [5]-[7]. Two approaches have been taken for implementing quadratic constraints in LMS updating. In the fixed loading approach, a fixed level of loading is chosen and added to the instantaneous estimate of the correlation matrix, $\mathbf{z}(k)\mathbf{z}(k)'$, at each step in the adaptation [1], [8]. The equivalent PLIC implementation is

$$e_p(k) = y_c(k) - \mathbf{w}_a(k-1)^H\mathbf{z}(k) \tag{14}$$
$$\mathbf{w}_a(k) = (\mathbf{I} - \alpha\lambda)\mathbf{w}_a(k-1) + \alpha\mathbf{z}(k)e_p^*(k) \tag{15}$$

In the scaled projection approach, the quadratic constraint is imposed on the adaptive weight vector at each update [6], [8]. This is accomplished by simply scaling back the norm of the adaptive weight vector whenever it exceeds the quadratic constraint, i.e.

$$e_p(k) = y_c(k) - \mathbf{w}_a(k-1)^H\mathbf{z}(k) \tag{16}$$
$$\tilde{\mathbf{w}}_a(k) = \mathbf{w}_a(k-1) + \alpha\mathbf{z}(k)e_p^*(k) \tag{17}$$
$$\text{if } ||\tilde{\mathbf{w}}_a(k)||^2 \leq \beta^2 \quad \mathbf{w}_a(k) = \tilde{\mathbf{w}}_a(k) \tag{18}$$
$$\text{if } ||\tilde{\mathbf{w}}_a(k)||^2 > \beta^2 \quad \mathbf{w}_a(k) = \tilde{\mathbf{w}}_a(k)\frac{\beta}{||\tilde{\mathbf{w}}_a(k)||}. \tag{19}$$

In RLS updating, the inverse of $\mathbf{R}_x$ or $\mathbf{R}_z$ is updated at each iteration, and there is no convenient method for incorporating diagonal loading into this term directly. LMS adaptation is computationally efficient but can be slow to converge. RLS converges quickly but more computationally complex. It is also more sensitive to mismatch than LMS.

6. VARIABLE LOADING RLS

A variable loading technique for RLS updating was proposed in [9] and is summarized here. Let $\tilde{\mathbf{w}}_a$ denote the standard MOE adaptive weight vector without quadratic constraints given by (12) and $\mathbf{w}_a$ denote the quadratically constrained weight vector (13). Rearranging (13), we have

$$\mathbf{w}_a = (\mathbf{I} + \lambda\mathbf{R}_z^{-1})^{-1}\mathbf{R}_z^{-1}\mathbf{p}_z = (\mathbf{I} + \lambda\mathbf{R}_z^{-1})^{-1}\tilde{\mathbf{w}}_a, \tag{20}$$

For small λ, $\mathbf{w}_a$ can be approximated using the first two terms of its Taylor series expansion about $\lambda = 0$,

$$\mathbf{w}_a \approx (\mathbf{I} - \lambda\mathbf{R}_z^{-1})\tilde{\mathbf{w}}_a \tag{21}$$
$$= \tilde{\mathbf{w}}_a - \lambda\mathbf{v}_a, \tag{22}$$

where $\mathbf{v}_a$ is defined as

$$\mathbf{v}_a = \mathbf{R}_z^{-1}\tilde{\mathbf{w}}_a. \tag{23}$$

If $\tilde{\mathbf{w}}_a$ does not satisfy the quadratic constraint, we can solve for λ by plugging (22) into the quadratic inequality constraint. This yields a second order polynomial in λ, with a closed form expression for the roots. When $||\tilde{\mathbf{w}}_a||^2 > \beta^2$, the roots of the polynomial are either two positive real values or a conjugate pair whose real part is positive. We would choose λ to be the smaller value in the first case, or the real part in the second case. When the real part is used, the constraint is not met but $||\mathbf{w}_a||^2$ is minimized.

The form of the approximate $\mathbf{w}_a$ in (22) provides a means for incorporating diagonal loading with RLS updating. At each recursion in the standard PLIC/GSC implementation, we have an estimate of $\mathbf{R}_z^{-1}$, denoted by $\mathbf{P}_z(k)$, as well as an estimate of the unconstrained weight vector, $\tilde{\mathbf{w}}_a(k)$. The current estimate for the constrained weight vector, $\mathbf{w}_a(k)$, can be obtained by substituting $\mathbf{P}_z(n)$ and $\tilde{\mathbf{w}}_a(k)$ into (22) and (23), solving for $\lambda(k)$, and substituting the result back into (22). The variable loading RLS algorithm (RLS-VL) iterates as follows:

$$\mathbf{g}(k) = \frac{\mu^{-1}\mathbf{P}(k-1)\mathbf{z}(k)}{1 + \mu^{-1}\mathbf{z}^H(k)\mathbf{P}(k-1)\mathbf{z}(k)}$$
$$\mathbf{P}(k) = \mu^{-1}\mathbf{P}(k-1)$$
$$\quad - \mu^{-1}\mathbf{g}(k)\mathbf{z}^H(k)\mathbf{P}(k-1)$$
$$e_p(k) = y_c(k) - \mathbf{w}_a^H(k-1)\mathbf{z}(k)$$
$$\tilde{\mathbf{w}}_a(k) = \mathbf{w}_a(k-1) + \mathbf{g}(k)e_p^*(k)$$

$$\text{if } ||\tilde{\mathbf{w}}_a(k)||^2 \leq \beta^2 \qquad \mathbf{w}_a(k) = \tilde{\mathbf{w}}_a(k) \qquad (24)$$

$$\text{if } ||\tilde{\mathbf{w}}_a(k)||^2 > \beta^2 \qquad \mathbf{v}_a(k) = \mathbf{P}(k)\tilde{\mathbf{w}}_a(k) \qquad (25)$$

$$a = ||\mathbf{v}_a(k)||^2 \qquad (26)$$

$$b = -2\Re e\left\{\mathbf{v}_a(k)'\tilde{\mathbf{w}}_a(k)\right\} \qquad (27)$$

$$c = ||\tilde{\mathbf{w}}_a(k)||^2 - \alpha^2 \qquad (28)$$

$$\lambda(k) = \frac{-b - \Re e\left\{\sqrt{b^2 - 4ac}\right\}}{2a} \qquad (29)$$

$$\mathbf{w}_a(k) = \tilde{\mathbf{w}}_a(k) - \lambda(k)\mathbf{v}_a(k). \qquad (30)$$

The variable loading technique does not contribute significantly to the overall computational load compared with the mainstream computations in the standard RLS updating.

7. SIMULATIONS

Application of the variable loading technique to the narrowband beamforming problem is described in [9]. Simulation results indicate that the RLS-VL technique provides superior beamforming results to standard RLS and RLS-SP in mismatch situations.

Simulation results for the DS-CDMA detection problem are shown in Figures 1-3 for a problem similar to those discussed in [1], [5] and [6].

In this example there are initially 7 users, each with SNR = 15 dB. The desired user has a L=31 Gold code spreading sequence, and the other spreading sequences are generated randomly. The desired signal experiences multipath propagation, resulting in a mismatch in the signature vector. The direct path has 90% of the signal energy. The second path ha s10% of the signal energy and is delayed by three chips. At $k = 500$ a new user is added with 25 dB SNR. Four techniques are considered: LMS, LMS with scaling (LMS-SP), RLS, and RLS with variable loading (RLS-VL). In Figure 1, the output SINR averaged over 20 trials is plotted for the first 1000 symbols. We see that the RLS algorithms adapt more quickly than the LMS algorithms initially and when the new user is added. Without the robustness constraint, RLS is also quick to suppress the desired user because of the mismatch. In Figures 2 and 3, the output SINR and weight vector norm are plotted for each technique for the first 9000 symbols. It can be seen that loss in SINR is directly related to the weight vector norm. Without the quadratic constraint, LMS eventually begins to degrade. Here the slow rate of convergence is beneficial. In Figure 4, the bit error rate (BER) is plotted vs. SNR of the desired user. Seven interferers have the same SNR as the desired user, and one interferer is 10 dB higher. All of the algorithms perform about the same for SNR below about 6 dB. After that, RLS and LMS begin to degrade because of the desired user suppression caused by the multipath component. The robust algorithms perform better, with RLS-VL having the best overall performance. It is also the most complex of the algorithms.
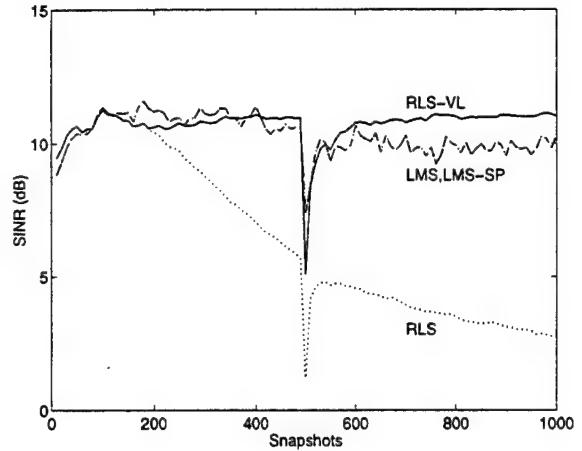

Figure 1: Output SINR for 1000 symbols
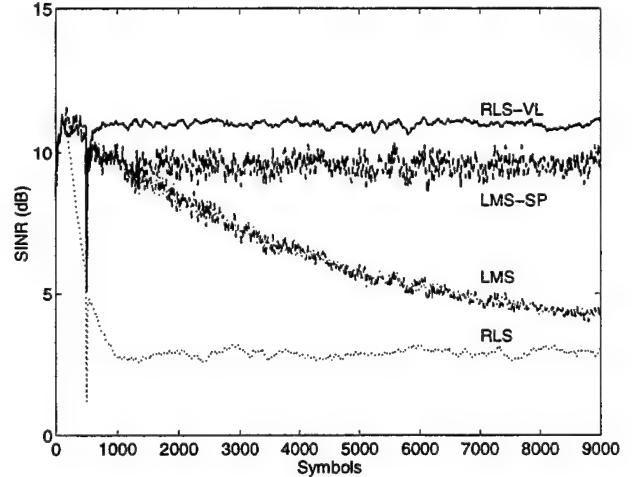
Figure 2: Output SINR for 9000 symbols
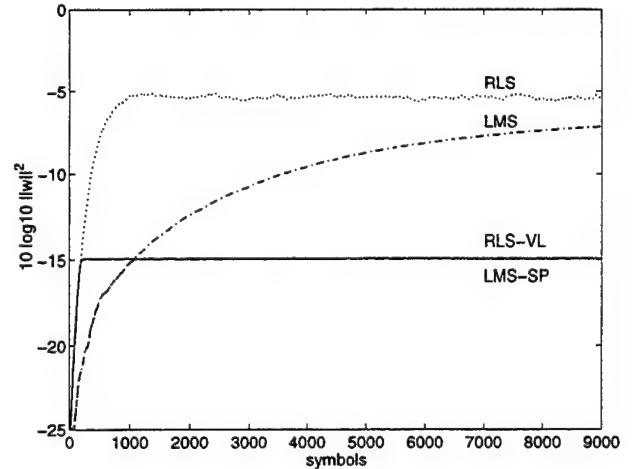
Figure 3: Square of weight vector norm

54

Figure 4: Bit Error Rate

8. CONCLUSION

A robust adaptive multi-user detector is presented in this paper. It uses variable loading to implement a quadratic inequality constraint with recursive least squares updating. This diagonal loading technique does not contribute significantly to the overall computational load. Simulations show that this technique offers better robust control over mismatch in modeling errors than other LMS and RLS adaptive techniques.

9. REFERENCES

[1] M. Honig, U. Madhow, and S. Verdu, "Blind Adaptive Multiuser Detection," *IEEE Trans. Inform. Theory*, vol. 41, no. 4, pp. 944-960, July 1995.

[2] B. D. Van Veen and K. M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE Acoust., Speech, Sig. Process. Magazine*, pp. 4-24, Apr. 1988.

[3] H. L . Van Trees, "Class notes: INFT836: 'Array Processing: Detection and Estimation Theory, Vol. IV'," April, 1998.

[4] L. J. Griffiths and C. W. Jim, "An Alternative Approach to Linearly Constrained Adaptive Beamforming," *IEEE Trans. Antennas Propagat.*, vol. 30, no. 1, pp. 27-34, Jan. 1982.

[5] J. B. Schodorf and D. B. Williams, "A Constrained Optimization Approach to Multiuser Detection," *IEEE Trans. Sig. Proc.*, vol. 45, no. 1, pp. 258-262, Jan. 1997.

[6] J. B. Schodorf and D. B. Williams, "Array Processing Techniques for Multiuser Detection," *IEEE Trans. Commun.*, vol. 45, no. 11, pp. 1375-1378, Nov. 1997.

[7] H. V. Poor and X. Wang, "Code-Aided Interference Suppression for DS/CDMA Communications-Part II: Parallel Blind Adaptive Implementations," *IEEE Trans. Commun.*, vol. 45, no. 9, pp. 1112-1122, Sep. 1997.

[8] H.Cox, R. Zeskind, M. Owen, "Robust Adaptive Beamforming," *IEEE Trans. ASSP.*, vol. 35, no. 10, pp. 1365-1376, Oct. 1987.

[9] Z. Tian, K. L. Bell, and H. L. Van Trees, "A RLS Implementation for Adaptive Beamforming Under Quadratic Constraint," *9th IEEE Workshop on Stat. Sig. and Array Processing*, Portland, Oregon, Sept. 1998.

ESTIMATION OF A JAMMED MULTIPATH CHANNEL DELAY PROFILE USING SUPERRESOLUTION METHODS AND A RANDOM PATH DECORRELATION TECHNIQUE

Frantz Bouchereau, David Brady,

Northeastern University
360 Huntington Av.
Boston MA, 02115

Colin Lanzl

Pinpoint Corporation
1 Oak Park
Bedford MA, 01730

ABSTRACT

In this paper, we extend a superresolution Pseudo-Noise sequence correlation method (SPM) to the time delay profile estimation of a jammed multipath channel. SPM is based on the MUltiple SIgnal Classification (MUSIC) algorithm which assumes a decorrelated signal covariance matrix, therefore, successful decorrelation of the inherently coherent multipath echos becomes a critical issue. Previously proposed path decorrelation techniques will fail in certain multipath environments and will be, in many cases, computationally expensive. We propose a decorrelation technique less likely to fail that will, in general, need fewer computations. We also propose a method that shortens the necessary snapshot transmission period and reduces the risk of channel variations during signal acquisition. Finally we show and analyze the results of underwater acoustic transmission experiments to which SPM was applied.

1. INTRODUCTION

In recent years, source localization applications have gained great importance. With the increase in computational power, we can now have real time systems that localize small active sources in a delimited area. Some of these systems use as a primary tool the estimation of the first time of arrival of multiple overlapping echos of a signal transmitted in a multipath channel to estimate the distance between the source and the receiving antenna [1]. It is evident, in this case, that the resolution capabilities of the channel delay profile estimator are a bottle-neck for accurate measurement of the first time

of arrival, and in consequence, a bottle-neck for the accuracy of the source localization algorithm. A method to estimate the time of arrival of overlapping echos in a multipath channel transmission using a superresolution algorithm was first presented in [2]. Later on, Manabe *et al* [3] presented a new algorithm, called SPM by the authors, that combines a conventional method of delay profile estimation using Pseudo Noise (PN) sequence correlations, with the MUSIC algorithm [4]. This method relies on the successful decorrelation of the inherently coherent multipath signals. The path decorrelation technique proposed in [3]—we will call it Frequency Smoothing (FS)— will fail in certain situations and will be computationally expensive in many cases. We propose a new decorrelation technique— Random Frequency Smoothing (RFS)— that will avoid these failure situations and will, in general, require less computations. We also analyze SPM in the presence of narrow-band jammers and present a simple way to null their effects. Finally, we present results when SPM and RFS were applied to underwater transmissions in the ocean. A method to allow the transmission of a higher number of snapshots in a short time interval is introduced.

2. THE SPM ALGORITHM

SPM is a combination of a conventional method to measure multipath delay profiles of a channel using PN-sequence correlations, and a superresolution algorithm such as MUSIC. The conventional scheme is shown in Figure 1. A *Maximal-Length Shift Register Sequence* (m-sequence), is transmitted through a D-path channel. The received signal is crosscorrelated with the original transmitted m-sequence. Due to the sharp m-sequence autocorrelation function , $z(\zeta)$ will have sharp peaks at every time of arrival of a signal echo. However,
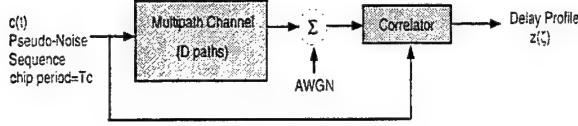
Figure 1: Conventional measurement of a multipath channel delay profile

the autocorrelation function of the m-sequence has a triangular spread of $\pm T_c$ seconds around the correlation peak, where T_c is the chip interval of the sequence. Then, echos that arrive with a time delay separation smaller that T_c will be difficult to resolve. The trivial solution is to increase the chip rate of the m-sequence to reduce the triangular spread. This becomes prohibitive in limited bandwidth and limited hardware complexity situations. An alternative solution is to use MUSIC to improve resolution. So let us consider transmitting an m-sequence through a channel with D discrete paths with different delay times $\{\tau_i \; ; \; i = 1, 2, ..., D\}$ which we want to estimate, and with an impulse response

$$h(t) = \sum_{i=1}^{D} h_i \delta(t - \tau_i) \; . \tag{1}$$

If the baseband modulation signal (an m-sequence) is $c(t)$ and ω_c is the RF carrier, then the baseband received signal is given by

$$v(t) = \sum_{i=1}^{D} h_i c(t - \tau_i) e^{j \omega_c \tau_i} + \eta(t) \; , \tag{2}$$

and the crosscorrelation of $v(t)$ with $c(t)$ is

$$z(\zeta) = \sum_{i=1}^{D} h_i e^{-j \omega_c \tau_i} r(\zeta - \tau_i) + \nu(\zeta) \; , \tag{3}$$

where $r(\zeta)$ is the m-sequence autocorrelation function, and $\nu(\zeta)$ is the crosscorrelation of the m-sequence and the noise process. If we sample the low resolution delay profile signal $z(\zeta)$ at M lags we obtain a delay profile vector $\mathbf{z} = [z(\zeta_1), z(\zeta_2), ..., z(\zeta_M)]^T$ given by

$$\mathbf{z} = \sum_{i=1}^{D} h_i e^{-j \omega_c \tau_i} \mathbf{r}(\tau_i) + \boldsymbol{\nu} \; , \tag{4}$$

where $\mathbf{r}(\tau)$ (the "steering vector") is defined as

$$\mathbf{r}(\tau) = [r(\zeta_1 - \tau), r(\zeta_2 - \tau), ..., r(\zeta_M - \tau)]^T \; , \tag{5}$$

$\mathbf{r}(\tau)$ can be calculated for any τ since the autocorrelation function $r(\cdot)$ is known. We define the "array manifold" matrix and the gain vector as

$$\boldsymbol{\Gamma} = [\mathbf{r}(\tau_1) \; \mathbf{r}(\tau_2) \; ... \; \mathbf{r}(\tau_D)] \; , \tag{6}$$

$$\mathbf{g} = [h_1 e^{-j \omega_c \tau_1} \; h_2 e^{-j \omega_c \tau_2} \; ... \; h_D e^{-j \omega_c \tau_D}]^T \; , \tag{7}$$

then we can write the delay profile vector in matrix notation

$$\mathbf{z} = \boldsymbol{\Gamma} \mathbf{g} + \boldsymbol{\nu} \; . \tag{8}$$

The covariance matrix of $\mathbf{z}$, which we shall consider as a function of the transmission carrier frequency for reasons that will become evident later, is given by

$$\mathbf{R}(\omega_{\mathbf{c}}) = \sum_{i,j} E\{h_i h_j^*\} e^{j \omega_c (\tau_j - \tau_i)} \mathbf{r}(\tau_i) \mathbf{r}(\tau_j)^H + E\{\boldsymbol{\nu} \boldsymbol{\nu}^H\} \; . \tag{9}$$

In matrix notation

$$\mathbf{R}(\omega_{\mathbf{c}}) = \boldsymbol{\Gamma} \mathbf{G}(\omega_{\mathbf{c}}) \boldsymbol{\Gamma}^H + \mathbf{R}_{\mathbf{noise}} \; . \tag{10}$$

Where $\mathbf{G}$ is the gain covariance matrix. The noise correlation matrix can be expressed as

$$E\{\boldsymbol{\nu} \boldsymbol{\nu}^H\} = \mathbf{R}_{\mathbf{noise}} = \sigma^2 \mathbf{R}_{\mathbf{0}} \; ; \; i = 1, 2, ..., M \; , \tag{11}$$

where $\mathbf{R}_{\mathbf{0}}$ is a symmetric matrix whose kl-th element is $r(\zeta_k - \zeta_l)$. Since we are interested in the eigendecomposition of $\mathbf{R}$, solving a generalized eigenvalue problem of the form

$$\mathbf{R} \mathbf{e}_i = \lambda_i \mathbf{R}_{\mathbf{0}} \mathbf{e}_i \; , \tag{12}$$

will be equivalent to whitening the noise. If $M > D$, and $\mathbf{G}$ is a positive definite matrix, and assuming for the moment that the echos are uncorrelated, then the signal covariance matrix $\mathbf{R} - \mathbf{R}_{\mathbf{noise}}$ will have rank D. Then it can be shown that $\mathbf{R}$ will have D generalized eigenvalues greater than the noise variance σ^2, and $M - D$ generalized eigenvalues equal to σ^2. The set of D eigenvectors $\{\mathbf{e}_i \; ; \; i = 1, ... D\}$ corresponding to the D largest generalized eigenvalues span the *signal subspace*, the set of $M - D$ eigenvectors $\{\mathbf{e}_i \; ; \; i = D + 1, ... M\}$, corresponding to the $M - D$ smallest generalized eigenvalues span the *noise subspace*. It can also be shown that the noise subspace is orthogonal to the columns of matrix $\boldsymbol{\Gamma}$, i.e., the noise subspace is orthogonal to the steering vectors evaluated at the true delays. Then, we can use a measure of orthogonality of all the possible steering vectors and the noise subspace eigenvectors. When the steering vector evaluated at a certain delay τ_{test} is orthogonal to the noise subspace, that τ_{test} is a true delay of the channel. Then we define the superresolution delay profile (SDP) as:

$$SDP(\tau) = \frac{\mathbf{r}(\tau)^H \mathbf{R}_{\mathbf{0}} \mathbf{r}(\tau)}{\sum_{i=D+1}^{M} [\mathbf{r}(\tau)^H \mathbf{e}_i]^2} \; . \tag{13}$$

This function will ideally peak to infinity at the D path delays.

In practice we use the *Minimum Description Length* criteria to estimate the number of paths D [5]. We repeat the transmission of $c(t)$ every T seconds to obtain N *snapshots* of the delay profile vector $\mathbf{z}$. We assume that after T seconds no more echos arrive, this implies that every transmission of $c(t)$ has to be followed by a silent period to allow all the echos to die out at the receiver. We also assume that through the N snapshots transmission, the channel is invariant. Then we estimate the covariance matrix as

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{j=1}^{N} \mathbf{z}_j \mathbf{z}_j^H . \qquad (14)$$

3. RANDOM FREQUENCY SMOOTHING

In the previous section we assumed that the channel path gains were decorrelated, however, in many multipath environments of interest this is not true. When we have a strong path coherency the rank of matrix $\mathbf{R} - \mathbf{R_{noise}} = \mathbf{\Gamma G \Gamma}^H$ is reduced to one. This rank deficient situation makes the correct separation of the signal and noise subspace impossible, consequently, SPM can not be applied unless we restore the signal covariance matrix rank to D. With uncorrelated paths, that is, when $Cov\{h_i, h_j\} = 0$, $i \neq j$, equation (9) becomes

$$\mathbf{R_{ideal}} = \sum_i |h_i|^2 \mathbf{r}(\tau_i) \mathbf{r}(\tau_i)^H + \mathbf{R_{noise}} . \qquad (15)$$

To increase the rank of matrix $\mathbf{\Gamma G \Gamma}^H$ to D, we use a technique proposed in [3] that we have named Frequency Smoothing (FS) because of its close relation to the array processing "spatial smoothing" algorithm [6]. FS consists in transmitting the m-sequence $c(t)$ at different uniformly spaced carrier frequencies around a central frequency ω_o, we then estimate K covariance matrices at each frequency and average to obtain the "smoothed" covariance matrix

$$\overline{\mathbf{R}}_{\mathbf{FS}} = \frac{1}{K} \sum_{k=1}^{K} \hat{\mathbf{R}}(\omega_k) . \qquad (16)$$

The carrier frequencies set is given by $\{\omega_k\} = \{\omega_0;\ \omega_0 + \Delta\omega;\ \omega_0 + 2\Delta\omega;\ ...;\ \omega_0 + (K-1)\Delta\omega\}$.

To measure the decorrelation performance of FS we use the following metric

$$m_{FS}(\Delta\omega) = max|EIG(\mathbf{R_{ideal}} - \overline{\mathbf{R}}_{\mathbf{FS}}, \mathbf{R_{noise}})| , \qquad (17)$$

where $EIG(\mathbf{A}, \mathbf{B})$ denotes the generalized eigenvalues of matrices $\mathbf{A}$ and $\mathbf{B}$. This metric will go to zero whenever $\overline{\mathbf{R}}_{\mathbf{FS}} = \mathbf{R_{ideal}}$, and will have maximum values when the two matrices differ the most. Clearly, the

choice of $\Delta\omega$, and K will affect the decorrelation performance. In [7] we show that for a two path channel with delay separation Δ_T, whenever $\Delta\omega = \frac{n\pi}{\Delta_T}$, for any even integer n, FS will fail to decorrelate the paths no matter how large K is made. So for an unknown channel, we can always choose a $\Delta\omega$ that will make the FS method fail.

As an example, Figure 2 shows two normalized (to a unity maximum correlation value) decorrelation metrics m_{FS} for a two path channel $(D = 2)$ with a path delay separation $\Delta_T = 20nS$; $K = 3$, and $K = 16$ carriers were used. It can be seen that this metric
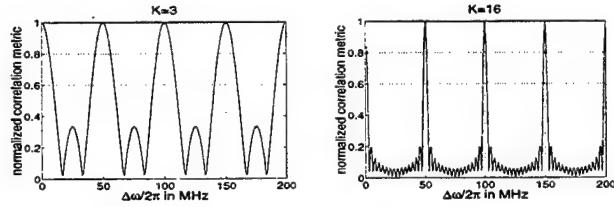
Figure 2: Decorrelation metric m_{FS} for $K = 3$, and $K = 16$ carriers

has peaks at the critical frequencies $\Delta\omega = 2\pi \frac{n}{2\Delta_T}$ for $n = 0, 2, 4, 6, 8$. As K increases the lobes of the correlation metric become narrower but the critical frequency values still show full correlation peaks. Observe that if we knew the channel, we could always look at the metric and choose a value of $\Delta\omega$ where the correlation value is minimum and with this, make the SPM algorithm perform at its best with a few number of carriers.

The importance of having a good decorrelation performance is illustrated in Figure 3. The plots show
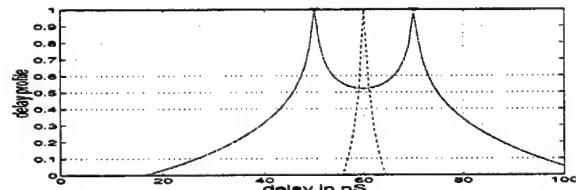
Figure 3: SDP $(K = 3$ FS) for a two path channel with worst case $\frac{\Delta\omega}{2\pi} = 50$ MHz (dashed plot), and with best case $\frac{\Delta\omega}{2\pi} = 17$ MHz (solid plot)

the superresolution delay profiles of the same two path channel that was used to find the metric m_{FS} in Figure 2; $K = 3$ carriers were used. The dashed plot shows the SDP when we used a critical $\frac{\Delta\omega}{2\pi} = 50$ MHz for the frequency smoothing, the solid plot is the SDP of the channel obtained with a best-case $\frac{\Delta\omega}{2\pi} = 17$ MHz which corresponds to a minimum value on the m_{FS} curve. Clearly, when the best-case $\Delta\omega$ is used, the two paths are resolved, when the critical worst-case $\Delta\omega$ is

used, the SPM algorithm fails and the two paths can not be resolved.

Since FS fails when we choose a certain $\Delta\omega$ or values close to it, we need to use a large number of carrier frequencies to make the correlation lobes of m_{FS} as narrow as possible. So even though using a large K will almost always result in a good path decorrelation, it may not be computationally feasible for implementation.

To avoid choosing a frequency value that will make the SPM algorithm fail, we propose transmitting the PN-sequences using a random, iid sequence of carrier frequencies. We call this technique random frequency smoothing (RFS). The new set of carriers becomes $\{\omega_k\} = \{\omega_0; \omega_0 + \Delta\omega_1; ...; \omega_0 + \Delta\omega_{K-1}\}$, where $\Delta\omega_i$ is a uniform random variable in the interval $[0, \Delta\omega_{max}]$. Using this random frequency grid, we will avoid the critical frequency values at least once with high probability. So a small number of carriers can be used since now making the lobes of the correlation metric narrower is not as important as before. So the number of computations can now be kept at a reasonable level.

4. SPM IN THE PRESENCE OF JAMMERS

In many real applications, the delay profile estimation of a multipath channel has to be done in the presence of jammers such as narrow band tones coming from other radio transmissions, or from the receiver and transmitter hardware. The SPM algorithm fails when a jammer is present in the channel and a simple extra step has to be taken to solve this problem. Below, we explain why SPM fails in the presence of a narrow band tone jammer $J = A_J e^{j2\pi f_J t}$.

Consider the jammed channel as the superposition of two linear time invariant systems with impulse responses $h_1(t)$ (multipath), and $h_2(t)$ (jammer). The total output containing the jammer and channel effects is the sum of the two system outputs. The impulse response of the first system is simply the multipath channel impulse responses given in equation (1). The impulse response of the second system must be one that produces the narrow band jammer J at its output. Such and impulse response has the form

$$h_2(t) = \frac{A_J}{|C(f_J)|} e^{j(2\pi f_J t - \angle C(f_J))} , \qquad (18)$$

where $C(f)$ is the Fourier transform of $c(t)$. The magnitude of $h_1(t)$ consists of D impulses at the time-delays of the channel, the magnitude of $h_2(t)$ is continuous in time with an amplitude of $\frac{A_J}{|C(f_J)|}$. The total jammer plus channel response will then be continuous at all time. So SPM is "observing" an infinite number of paths and will fail to resolve them because the number of samples M will not be large enough.

A quite easy way to null the jammer effects is to consider jammers and noise as one colored noise with unknown distribution. Then, we measure L noise snapshots at the output of the correlator and estimate a noise covariance matrix before we start transmitting the m-sequences. The generalized eigenvalue problem needed to find the noise and signal subspaces, is solved using this estimated noise covariance matrix. This will effectively null the colored noise and jammer effects. This method works not only for narrow band tone jammers but for any type of interference that remains stationary during the noise covariance matrix and signal covariance matrix estimation periods.

5. REAL DATA ANALYSIS

Underwater transmission experiments in the ocean were carried out to test the performance of SPM and RFS.

5.1. Signal transmission scheme

The cascade transmitter-receiver response was measured to have a 3 dB bandwidth of 10 KHz centered at 35 KHz. So the m-sequence chip rate was set to $\frac{1}{T_c} = 4$ KHz ($T_c = 0.25\ mS$), to allow a $\Delta\omega_{max} = 2\pi\cdot(1\ KHz)$ variation of the RFS carriers.

As stated at the end of section 2, to estimate the sample covariance matrix $\hat{R}$, N signal snapshots followed by silent periods longer than the delay spread of the channel are transmitted. To avoid the loss of signal snapshot transmission time due to these silent periods, we propose transmitting back-to-back preferred m-sequences. A member belonging to a family of n preferred m-sequences (n=6 for this experiment) has a thumb-tack shaped autocorrelation, and a low crosscorrelation peak magnitude with every other member in the family. The receiver will now consist of a bank of n correlators, one for each preferred m-sequence. Then, the silent period of sequence 1 will be given by the low crosscorrelation levels of the following $n-1$ m-sequences. The nonzero crosscorrelations will increase the noise level at the correlator output, but this noise level will still be acceptable for good performance. On the other hand, we will be able to increase the number of sequence transmissions during the channel coherence time and the signal covariance matrix $\mathbf{R}$ will be estimated more accurately.

Figure 4 shows the signal scheme used for the experiment. Three 8 KHz chip rate Barker sequences were transmitted at three different times to observe any possible channel variations, also, we will compare the SPM
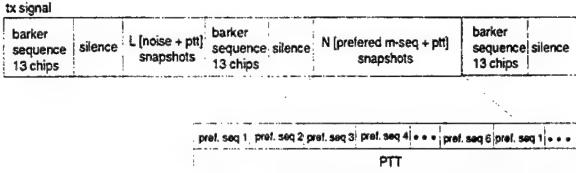
Figure 4: Signal transmission scheme for the experiment

superresolution delay profile estimates with the conventional PN-correlation method of Figure 1 using these Barker Sequences as the pseudo noise signals. We will also compare SPM to a recursive least squares (RLS) channel estimation algorithm with 100 taps each one separated by $\frac{T_c}{2}$ seconds. The Barker sequence chip rate was set to twice the preferred m-sequences rate to justify that when using a superresolution method we can obtain equal or better resolution using half of the transmission bandwidth.

In order to track phase rotations due to water and hydrophone movements, a Phase Tracking Tone (PTT) was transmitted on top of the preferred m-sequence snapshots. We will treat this tone as a jammer when running SPM, for this reason, the PTT was also transmitted during the L noise snapshots measurement period. Figure 5 shows the frequency spectrum of a received snapshot in the pond. It is clear that the spec-
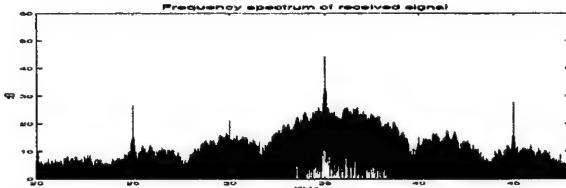

Figure 5: Frequency spectrum of a received snapshot

trum is not centered at 35 KHz but has a shift of a random $\Delta\omega$. Note the strong spectral lines at 25, 30, and 45 KHz, these are jammers caused possibly by the hardware and/or by other acoustic signals present in the channel, we also see the 35 KHz PTT artificially placed to track the signal phase. The effects of these jammers will be included in the estimated noise covariance matrix and will be nulled when solving the generalized eigenvalue problem of equation (12).

5.2. Practical issues

Many factors affect the SPM algorithm, we list them here and discuss them in detail in [7].

1. *Limited Transmission time*: Transmission time is either limited by the stationarity period of the

channel, or by the acquisition hardware. If we do not use a back-to-back preferred m-sequence transmission scheme, then the number of possible snapshot transmissions can be very low and the SDP will have bias errors and low resolution.

2. *Large multipath delay spread*: To sample a long channel delay spread, we need large number of samples M. If we do not have enough number of snapshots to satisfy $N > M$ we will not be able to sample the complete delay spread. This causes errors on the SPM estimates since some echos are sampled only for a fraction of their duration. Also, note that a large M increases the computational load.

3. *Limited sampling rate at correlator output*: For reasons mentioned in the previous points, the sampling rate at the correlator output needs to be reduced in order to keep a low computational complexity, and/or to make sure that $N > M$. With low sampling rates, we will observe spurious peaks that can be confused as a true path delay estimate.

4. *Estimation of the number of paths D*: Overestimating the number of paths causes spurious peaks on the SPM estimates since we are not using all the noise subspace eigenvectors in the denominator of equation (13), then, not all the spurious peaks from each individual eigenvector are being averaged to zero. On the other hand, underestimating the number of paths leaves us with low resolution estimates with a very low difference between the SDP floor and its peaks.

5. *Hardware nonlinearities and limited resolution*: The data was transmitted using a push-pull amplifier. This causes quantization noise. Another source of distortion will be the limited 8 bit resolution of the analog to digital converter.

5.3. Experimental results

We now present the superresolution delay profile estimates obtained in ocean transmissions. The measured signal to noise ratio at the input of the correlator was 22 dB. $K = 10$, and $K = 20$ carriers were used for RFS. Using MDL, the estimated number of paths was $\hat{D} = 60$. The back-to-back preferred m-sequence scheme allowed the transmission of $N = 170$ snapshots to estimate the covariance matrix. Figures 6, and 7 shows the channel SDP estimates compared with the double bandwidth Barker crosscorrelations and with the RLS channel estimates. All the estimate plots were

shifted to start at $t = 0$, and only 5 mS of the total 28 mS measured delay spread of the channel are presented for visual convenience. We observe that with
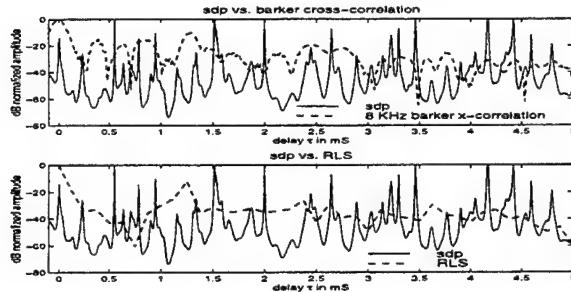

Figure 6: SDP estimate with K=10 compared to a Barker crosscorrelation and RLS
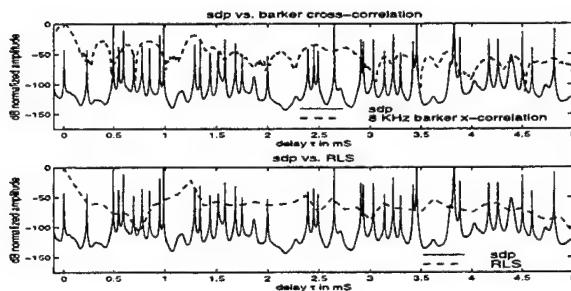
Figure 7: SDP estimate with K=20 compared to a Barker crosscorrelation and RLS

$K = 20$ the peaks of the SDP estimate are sharper due to a better path decorrelation. The SDP estimates have more than one peak inside most of the RLS and barker crosscorrelation peaks, showing that SPM has a better resolution than the double bandwidth barker crosscorrelation, and the $\frac{T_c}{2}$ RLS. The maximum resolution observed in the SDP plots is 0.06 mS which is equivalent to $0.24T_c$. We conclude that we can obtain a much higher resolution by using SPM than by increasing the bandwidth of the m-sequences to reduce T_c.

6. CONCLUSIONS

A method to estimate a multipath channel delay profile with high resolution was described. SPM does not work well when the signals coming into the sensors are correlated. In many multipath channels of interest, the signal echos are completely correlated and a decorrelation processing has to be implemented. A new low complexity random frequency smoothing technique was presented. This technique allows the path decorrelation with a small number of carriers making

its implementation feasible. Narrow band tone jammers present in the channel make SPM fail, and we described a way to null this jammer effects. Experiments were conducted to show how the SPM-RFS algorithms perform with hardware nonlinearities and limitations, and when jammers are present in the channel. We have also presented a transmission scheme in which no silent periods are needed between each signal snapshot. We found that SPM worked well in practice and performed better than other estimation methods that used twice the signal bandwidth. We conclude that SPM implementation in real systems is feasible, and allows the estimation of high resolution channel delay profiles using a moderate signal bandwidth (smaller than the channel bandwidth).

7. REFERENCES

[1] Jay Werb, Colin Lanzl, "Designing a positioning system for finding things and people indoors", *IEEE Spectrum Magazine*, September 1998, pp. 71–78.

[2] Alfred M. Bruckstein, Tie-Jun Shan, Thomas Kailath, "The Resolution of Overlapping Echos", *IEEE Trans. Acoustics, Speech, and Signal Proc.*, vol. ASSP-33, No. 6, December 1985, pp. 1357–1367.

[3] Takeshi Manabe, Hitoshi Takai, "Superresolution of Multipath Delay Profiles Measured by PN Correlation Method", *IEEE Trans. Antennas and Prop.*, vol. 40, No. 5, May 1992, pp. 500–509.

[4] R.O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Trans. Antennas and Prop.*, vol. AP-34, March 1986, pp. 276–280.

[5] Mati Wax, T. Kailath, "Detection of Signals by Information Theoretic Criteria", *IEEE Trans. Acoustics, Speech, and Signal Proc.*, vol. ASSP-33, No. 2, April 1985, pp. 387–392.

[6] Tie-Jun Shan, Mati Wax, T. Kailath, "On Spatial Smoothing for Direction Of Arrival Estimation of Coherent Signals", *IEEE Trans. Acoustics, Speech, and Signal Proc.*, vol. ASSP-33, No. 4, August 1985, pp. 806–811.

[7] Frantz Bouchereau, David Brady, "Estimation of Multipath Delay Profiles in Underwater Acoustic Communication Experiments using Superresolution Methods", in preparation, contact first author at fboucher@cdsp.neu.edu for a preprint.

NULLING OVER EXTREMELY WIDE BANDWIDTHS WHEN USING STRETCH PROCESSING

Richard M. Davis, Jose A. Torres, J. David R. Kramer, Ronald L. Fante

The MITRE Corporation
202 Burlington Road
Bedford, MA 01730-1420
rmdavis@mitre.org, torres@mitre.org, dkramer@mitre.org, rfante@mitre.org

ABSTRACT

This paper describes a technique for performing adaptive jammer nulling over extremely wide bandwidths that is applicable to radar systems which use linear frequency modulated (LFM) waveforms and stretch processing. It is assumed that the range uncertainty of the target is a small percentage of the equivalent range extent of the uncompressed pulse. The assumptions allow the cancellation to take place in either the time domain using a narrowband sliding filter that keeps up with the chirp rate or in the frequency domain. The new approach supports nulling performance over a gigahertz of bandwidth comparable to that previously achieved over a few megahertz using approximately the same number of degrees of freedom.

1. BACKGROUND

Traditional sidelobe canceller architectures obtain samples of the jamming signals from spatially separated auxiliary antennas. The samples are multiplied by complex weights and then summed and combined with the output of the main antenna to minimize the interference. It takes one auxiliary antenna and one adaptive weight to null a single narrowband sidelobe jammer. As the bandwidth of the jammer increases, the spectrum spreads through the sidelobes of the pattern so that each frequency is received with a slightly different gain. The number of sidelobes which the jammer spreads through is equal to the product of the difference in arrival time of the jammer plane wave across the face of the antenna with the jammer bandwidth. The product is referred to as the time-bandwidth or TB product. The TB product of large modern microwave radars having ten or twenty percent tunable bandwidths can easily exceed ten.

A common rule of thumb for traditional spatial cancellers is that it takes two adaptive weights (two auxiliary antennas) to cancel a jammer which spreads through a single sidelobe. Thus, it could take twenty adaptive weights to null a single jammer having a TB product of ten using the traditional architecture. Since each adaptive channel requires a separate receiver that must operate instantaneously over the entire cancellation bandwidth, the traditional wideband adaptive spatial processor quickly becomes cumbersome and costly.

The number of auxiliary antennas, or equivalently the number of independent spatial samples, can be reduced by using adaptive finite impulse response (FIR) filters (space-time processing) or by dividing the bandwidth into subbands and operating an independent canceller in each. The use of adaptive FIR filters or subbanding, however, cannot reduce the number of required adaptive weights.

2. POST-STRETCH PROCESSOR

In this section, we introduce an alternative to the traditional approaches for nulling over extremely wide bandwidths. The technique only works for LFM waveforms and stretch processing. Stretch processing is a technique which enables the use of wideband waveforms with narrowband processing. Stretch processing is most effective when the range uncertainty is a small fraction of the range extent of the uncompressed pulse. It entails putting a deramping oscillator into the first mixer to translate each target return into a tone which identifies its range (see Figure 1). The processing maps range into frequency. Detection can be implemented in the frequency domain by taking an FFT of the deramped signal.

The minimum bandwidth which must be carried after deramping is given by

$$B_{MIN} = \left(\frac{B}{T_u}\right)\frac{2\Delta R}{c},\qquad(1)$$

where

B = signal (chirp) bandwidth,
T_u = uncompressed pulse width,
ΔR = range uncertainty = range window, and
c = speed of light.

Let us assume that a sidelobe jammer is present which spreads its energy uniformly over the chirped bandwidth. If we insert a baseband filter having bandwidth B_{MIN} in the receiver following the deramper, then at any instant of time, the output of the filter will display a slice of the jammer spectrum having width equal to B_{MIN}. The entire jammer spectrum will slide by the filter window as time runs from zero to T_u seconds (see Figure 2).

Two approaches for implementing the canceller in the time domain are possible: (a) the sliding window approach, and (b) the multiple fixed window approach. Cancellation in the time domain takes place after the deramper, but before the FFT (see Figure 1). As previously noted, the portion of the jammer

$s(t) = \exp(j2\pi f_d t)$
= deramping local
oscillator frequency

• Adaptive weights are time-varying
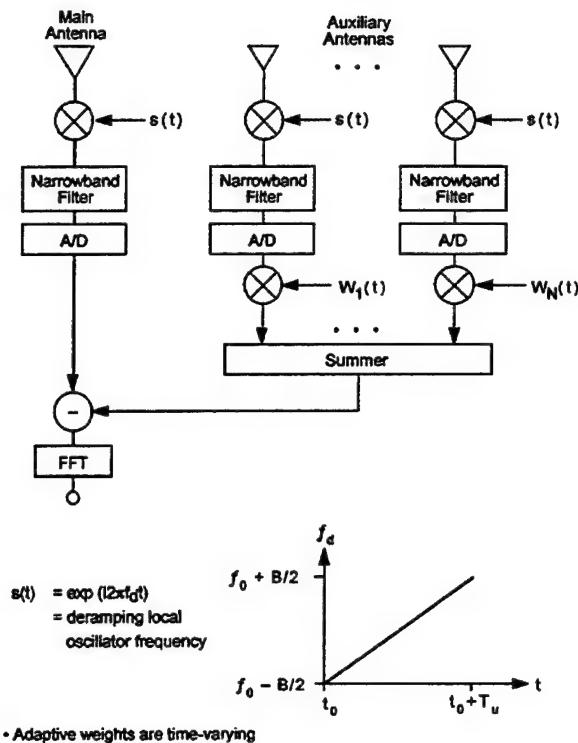
**Figure 1. Time Domain Post-Stretch Processing Canceller**
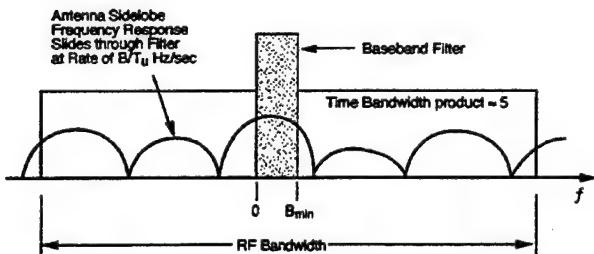
Figure 2. Frequency Response of Sidelobe Jammer Sliding Through Baseband Filter

spectrum that appears at the output of the baseband filter is constantly changing as the deramping frequency slides across the signal bandwidth. Using the sliding window approach, the adaptive weights must be changed fast enough to keep up with the changing filter response. A recursive algorithm such as the least mean square (LMS) algorithm is well suited for determining the time varying weights in the sliding window implementation.

The canceller could alternatively be implemented in the time domain using multiple non-overlapping fixed windows. The windows are obtained by segmenting the digitized samples out

of the baseband filter. All of the samples in each window in the main and auxiliary channel are stored and either all or some of them are used to calculate the adaptive weights. The weights are applied retroactively to all of the auxiliary samples in the window. The weighted auxiliary samples are summed and subtracted from the samples in the main channel to null the interference. An algorithm such as sample matrix inversion (SMI) could be used to calculate the adaptive weights when using fixed windows. It is noteworthy that the windows represent subbands that are generated automatically by the deramping and narrowband filtering operations.

The canceller can also be implemented in the frequency domain. The frequency domain approach requires that the deramped data in the main and in each auxiliary channel, which has been narrowband filtered and digitized, be transformed into the frequency domain using a FFT. After transformation, the main and each auxiliary will consist of hundreds to thousands of frequency bins. Only some of the frequency bins will correspond to the range window of interest, because the narrow-band filter is generally chosen to be wider than B_{MIN}. Since target returns will be above the thermal noise level after the FFT, we must either calculate the adaptive weights prior to transmission or use out-of-band correlation to prevent the canceller from nulling target returns. The out-of-band correlation technique entails using a small number of frequency bins in each auxiliary channel separated from the bins corresponding to the range window of interest to calculate the cross correlations needed to determine the weights. Out-of-band correlation can be implemented using bins on one or on both sides of the bin to be cancelled. The weighted bins are summed and subtracted from the main channel bin to be cancelled (see Figure 3).

$R(\omega)$ = Residue in frequency domain
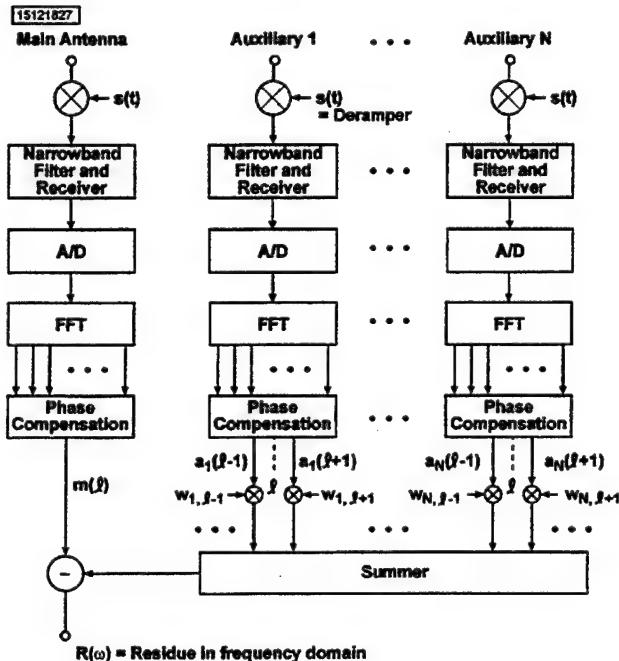
Figure 3. Frequency Domain Post Stretch Processing Canceller

3. ANALYSIS

The top insert in Figure 4 shows a block diagram of the architecture for the time domain canceller. The time-varying adaptive filter $g(t, t')$ is derived in Appendix A where it is shown that the transform of $g(t, t')$ is given by

$$G(t, \omega) = \gamma_0 \exp(i\omega t) H(2\beta t - \omega), \qquad (2a)$$

or

$$g(t, t') = \gamma_0 \int \exp(i\omega t) H(2\beta t - \omega)\exp(i\omega t')df, \qquad (2b)$$

where:

γ_0 = real constant,

β = chirp slope $= \pi B/T_u$,

$H(\omega)$ = Fourier transform of impulse response $h(t)$ of main antenna in direction of jammer, and

ω = $2\pi f$.

Equation 2 shows that the adaptive filter must match the time varying filter function in the main channel. The adaptive filter $G(t, \omega)$ is implemented by applying a set of complex weights to the auxiliary channels and then summing them together.

The bottom insert in Figure 4 shows a block diagram of the architecture for the frequency domain canceller. The functional form of the adaptive filter in the frequency domain is shown in Appendix B to be

$$G_0(\omega, \omega') = h((\omega - \omega')/2\beta), \qquad (3)$$
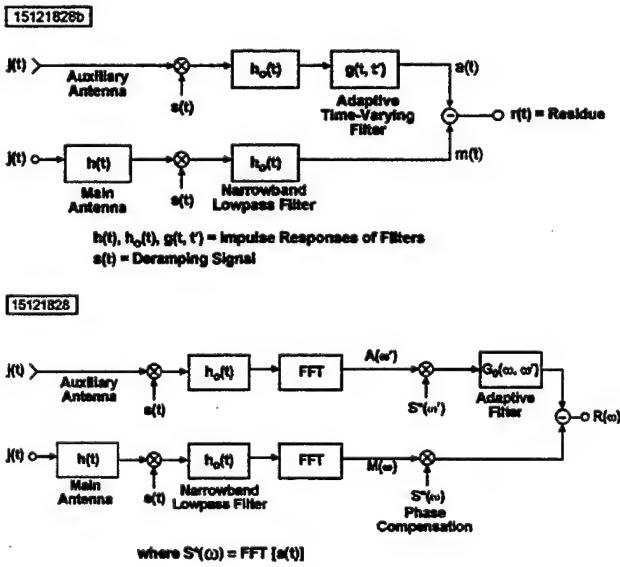
Figure 4. Adaptive Filters in Time (Top) and Frequency Domains (Bottom)

where:

ω, ω' = radian frequencies in main and auxiliary channels, respectively,

$h(t)$ = impulse response of main antenna in direction of jammer,

= $FFT^{-1}(H(\omega))$.

In order to make the frequency transfer function in the main channel equal to the composite transfer function of all the auxiliary channels, it was necessary to impose a phase compensation across the frequency bins equal to the Fourier Transformation of the deramping signal $s(t)$.

Given that ω is the center of the frequency bin in the main channel to be cancelled, (3) tells us that the adaptive weights applied to frequency bins in the auxiliary channels are only a function of the frequency separation $\omega-\omega'$. In other words, once we derive the weights to cancel one main channel frequency bin, we will be able to use those same weights (applied to different auxiliary frequency bins) to cancel the jamming in all other main channel bins. Equation (3) also tells us that the number of weights required to null the interference is determined by the effective length (the support) of the impulse response $h(\omega/2\beta)$. The support is typically small – about 5 to 10 frequency bins. Thus the jamming can be removed from the range bins of interest by sliding a fixed set of 5 to 10 weights across the bins in each of the auxiliaries and then summing the outputs and subtracting from the main channel bin, while keeping the separation between the auxiliary and main channel bins fixed.

We refer to the frequency bins as taps because of the analogy with a tapped delay line. In fact, because of the mapping between frequency and time created by the stretch processing, the frequency bins really are time taps. The observation explains why only one set of auxiliary channel weights is needed to cancel the jamming in all main channel frequency bins. The cross-correlations used to calculate the adaptive weights depend only upon the difference in arrival time of the jamming between the two time taps being cross-correlated. The frequency domain approach is analogous to space-time processing.

The adaptive weights in the frequency domain are given by the inverse of the covariance matrix (R) of the auxiliaries times the cross-correlation vector (P) between the main and auxiliaries. Individual elements of R and P are calculated using the functionality shown in Equations (4) and (5). Note that k can designate any one of the frequency domain taps as long as it is not one of the taps being cancelled (i.e., as long as it is out-of-band).

$$r\big(a_s(k), a_t(k+\ell)\big) = \frac{1}{M}\sum_{k=1}^{M} a_s(k)a_t^*(k+\ell), \qquad (4)$$

$$p\big(m(k), a_t(k+\ell)\big) = \frac{1}{M}\sum_{k=1}^{M} m(k)a_t^*(k+\ell), \qquad (5)$$

65

where:

$a_s(k)$ = signal on k^{th} frequency bin in s^{th} auxiliary channel,

$m(k)$ = signal on k^{th} frequency bin in main channel.

4. NUMERICAL RESULTS

The post-stretch processor was simulated on the computer in both the time and frequency domains. The radar waveform was modeled as a 100 μsec pulse chirped over 1,000 MHz of bandwidth at a center frequency of 7.0 GHz. After deramping, the bandwidth was reduced from 1,000 to 20 MHz. The bandwidth decimation resulted in a processing gain of 17 dB on the target return prior to analog-to-digital (A/D) conversion. The remainder of the 50 dB of pulse compression gain occurred in the FFT following A/D conversion. A total of 2,000 complex baseband samples were taken within the uncompressed pulse.

The simulation used a 264 element linear array divided into 12 subarrays of 22 elements each. Phase steering was used within a subarray and time steering between subarrays. The array was steered approximately 30 degrees off normal in all runs. The auxiliaries were chosen to be individual elements out of the array. The number of auxiliaries was varied from one to eight and were located at elements 1, 264, 97, 188, 102, 181, 241, and 260. The jammers were located between 34 and 44 degrees away from the steering direction which resulted in BT products that varied from 10 to 13.

Figure 5 compares the performance of time and frequency domain cancellers using the SMI algorithm to calculate the

15121838P

Eight Auxiliaries

Three Jammers

B_{LPF} = 20 MHz

2000 A/D Samples within T_u

Jammer Time-Bandwidth Product = 11.5; 10.0; 13.9

Avg # of A/D Samples per Sidelobe = 175; 200; 144

Maximum SIR = 18.5 dB

Ns = Time-Domain SMI Block Length

Nf = Number of Adaptive Frequency Bins

Two-Sided Out-of-Band Training in Freq-Domain with Phase Compensation (480 Frequency Samples)
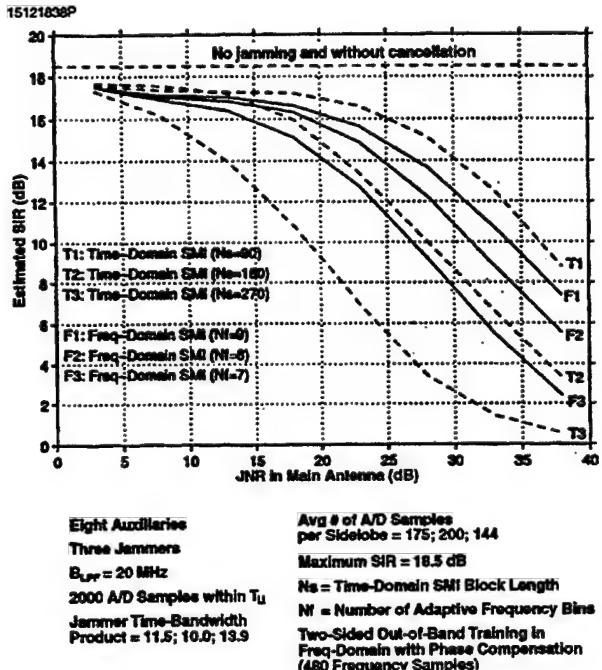
Figure 5. Comparison of Three Jammer Cancellation Performance in Time and Frequency Domains Using SMI Algorithm

weights. The 2,000 samples were divided into non-overlapping blocks and a separate canceller was operated within each block when modeling the time domain canceller. A number of sampling options was compared when forming the covariance matrix. Using all samples in a block provided the best performance. The figure compares the performance of block sizes consisting of 90, 180, and 270 samples. The canceller consisted of eight auxiliaries and there were three jammers in the threat. The number of samples per sidelobe was 175, 200, and 144, respectively for the three jammers. The figure plots signal-to-interference ratio (SIR) after pulse compression versus jammer-to-noise ratio (JNR) received in the main antenna. SIR is equivalent to dividing the signal-to-noise ratio (SNR) by the residue-to-noise-ratio after cancellation. SNR was 18.5 dB in the quiescent (no jamming) environment. All results represent the average of 50 Monte Carlo runs. Figure 5 also shows cancellation performance in the frequency domain using 7, 8, and 9 frequency bins (time taps) in each of the auxiliaries. The cancellation performance realized in the frequency domain is comparable to that obtained in the time domain.

Figures 6 and 7 show additional frequency domain results for three auxiliaries operating against a single jammer. The figures plot the output power in each of the 2,000 frequency bins. The range window of interest spans bins −50 to 50 and a target return is shown in bin zero. Figure 6 does not include the quadratic phase compensation while Figure 7 does. In general, the residue (within bins ±50) is lower with phase compensation than without it. Twelve frequency bins in each of the three auxiliaries near the bin to be cancelled in the main channel were weighted, summed, and subtracted from the bin in the main channel to null the interference. The process was repeated in each of the 2,000 bins using the same set of 36 weights. In the absence of phase compensation, the cancellation performance deteriorates as one moves away from the out-of-band bins used to train the weights. Although we have shown all 2,000 bins in the spectrum, the range window of interest will occupy only a small percentage of the entire spectrum. Thus depending upon the exact size of the range window of interest and the requirement of jammer nulling, phase compensation may not be needed.

5. SUMMARY

We have demonstrated the feasibility of nulling sidelobe jammers over extremely wide bandwidths on systems which use stretch processing. The proposed technique exploits the mapping between time and frequency implicit in stretch systems. It was shown that nulling could be performed in the time or frequency domains. Nulling in the time domain takes place at the output of a narrowband filter. The Jammer spectrum slides through the filter during the arrival time of the uncompressed pulse. The time domain approach is analogous to traditional cancellers which use subbanding while the frequency domain approach is analogous to space-time processing.

It was shown that the transfer function of the auxiliary channels can be made identical to that of the main antenna by applying a filter to the spectrums of the auxiliary channels
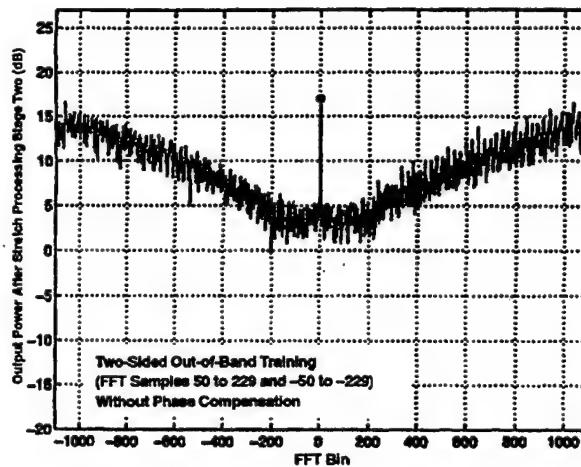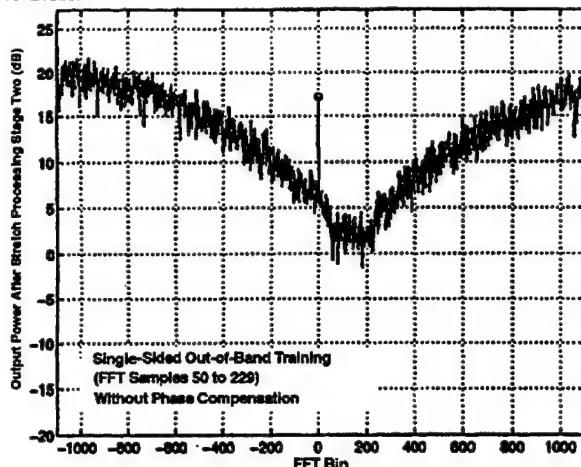
66

Single-Sided Out-of-Band Training
(FFT Samples 50 to 229)
Without Phase Compensation

Two-Sided Out-of-Band Training
(FFT Samples 50 to 229 and −50 to −229)
Without Phase Compensation

Three Auxiliaries	JNR in Main Antenna = 34 dB
Single Jammer	Maximum SIR = 18.5 dB
B$_{LPF}$ = 20 MHz	Target Located at FFT Bin 0
2000 A/D Samples within T$_u$	Number of Adaptive Freq Bins = 12
Jammer Time-Bandwidth Product = 11.5	Freq-Domain SMI Block Length = 180

Figure 6. Three Auxiliaries Operating Against One Jammer in Frequency Domain Without Phase Compensation

equal to the impulse response of the main antenna. The argument of the impulse response was shown to depend only upon the frequency difference between the main channel bin being cancelled and the auxiliary bin that is weighted and subtracted from the main channel bin. The observation showed that once a set of frequency domain weights is found that nulls a single frequency bin in the main channel, those same weights can be applied to other frequency bins in the auxiliaries to null any other bin in the main channel.
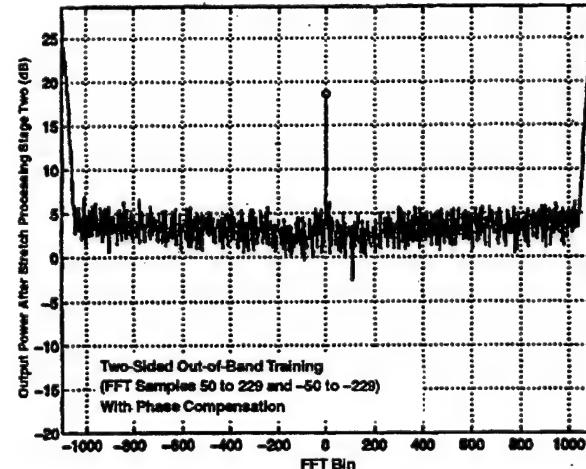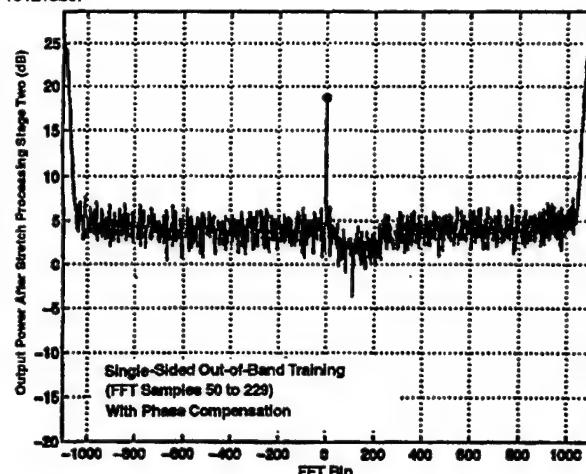
Single-Sided Out-of-Band Training
(FFT Samples 50 to 229)
With Phase Compensation

Two-Sided Out-of-Band Training
(FFT Samples 50 to 229 and −50 to −229)
With Phase Compensation

Three Auxiliaries	JNR in Main Antenna = 34 dB
Single Jammer	Maximum SIR = 18.5 dB
B$_{LPF}$ = 20 MHz	Target Located at FFT Bin 0
2000 A/D Samples within T$_u$	Number of Adaptive Freq Bins = 12
Jammer Time-Bandwidth Product = 11.5	Freq-Domain SMI Block Length = 180

Figure 7. Three Auxiliaries Operating Against One Jammer in Frequency Domain with Phase Compensation

6. APPENDIX A: DERIVATION OF TIME-DOMAIN ADAPTIVE FILTER

The voltages in the auxiliary (a) and main (b) channels shown in the top insert in Figure 4 can be expressed as

$$a(t) = \int dt' g(t, t') \int d\tau j(\tau) s(\tau) h_0(t' - \tau), \qquad (A-1)$$

$$m(t) = \int dt' h_0(t - t') s(t') \int d\tau j(\tau) h(t' - \tau). \qquad (A-2)$$

67

We require $a - m = 0$ for arbitrary $j(\tau)$

$$\int d\tau\, j(\tau)\left[s(\tau)\int dt'\, g(t,\, t')h_o(t'-\tau)\right.$$
$$\left. -\int dt'\, h_o(t-t')h(t'-\tau)s(t')\right] = 0. \qquad (A\text{-}3)$$

Therefore, for $s(t) = \exp\!\left(i\beta t^2\right)$, where $\beta = \dfrac{\pi B}{T_u}$ we require

$$\int dt'\, g(t,\, t')h_o(t'-\tau)$$
$$= \int dt'\, h_o(t-t')h(t'-\tau)\exp\!\left(i\beta\!\left(t'^2 - \tau^2\right)\right). \qquad (A\text{-}4)$$

If we write

$$g(t, t') = \int df\, G(t, f)\exp(i2\pi f t'), \qquad (A\text{-}5)$$

and substitute Equation (A-5) into Equation (A-4)

$$\int df\, G(t,\, f)\,H_o(-f)\exp(i2\pi f\tau)$$
$$= \int dt'\, h_o(t-t')h(t'-\tau)\exp\!\left(i\beta\!\left(t'^2 - \tau^2\right)\right), \qquad (A\text{-}6)$$

where $H_o(f)$ is Fourier Transform of $h_o(\tau)$. Now, take an inverse transform of Equation (A-6)

$$G(t,\, f) = \frac{1}{H_o(-f)}\int dt'\, h_o(t-t')\exp\!\left(i\beta t'^2\right)$$
$$\cdot\int d\tau\, h(t'-\tau)\,\exp\!\left(-i\beta\tau^2\right)\exp\left(-i2\pi f\tau\right). \qquad (A\text{-}7)$$

Now replace h_o and h by their Fourier transforms and perform the integrals over τ and t' to obtain

$$H_o(-f)G(t,\, f) = K^2\int d\xi\, d\eta\, H_o(\eta)H(\xi)\exp\!\left(i\omega^2/4\beta\right)$$
$$\cdot\exp\!\left(-i\eta^2/4\beta\right)\exp\!\left(i\xi(\eta+\omega)/2\beta\right)\exp(i\eta t). \qquad (A\text{-}8)$$

If $H_o(\eta)$ = constant over $\pm\eta_{MAX}$, and $z = \xi - 2\beta t$, then after considerable manipulation, it can be shown that

$$G(t,\, \omega) = \gamma_o'\exp\left(i\omega t\right)\int_{-\eta_{MAX}}^{\eta_{MAX}} H(z+2\beta t) = \exp\!\left(i(z+\omega)^2/4\beta\right)dz,$$
$$(A\text{-}9)$$

for $|\omega| < \eta_{MAX}$ and γ_o' a constant. Note that if $|\alpha| \gg \eta_{MAX}$, then by stationary phase, the integral vanishes because the stationary point is outside the range of integration. Applying the stationary phase approximation to the integral in Equation (A-9) yields

$$G(t,\, \omega) \approx \gamma_o\exp(i\omega t)H(2\beta t-\omega), \qquad (A\text{-}10)$$

for $|\omega| < \eta_{MAX}$.

7. APPENDIX B: DERIVATION OF FREQUENCY DOMAIN FILTER

Taking the Fourier Transform of Equations (A-1) and (A-2), we obtain

$$A(\omega) = \int d\omega'\, G(\omega,\, \omega')\int d\omega''\, J(\omega'')S(\omega'-\omega'')H_o(\omega'), \quad (B\text{-}1)$$

$$M(\omega) = H_o(\omega)\int d\omega''\, J(\omega'')H(\omega'')S(\omega-\omega''). \qquad (B\text{-}2)$$

Since we desire the residue $R(\omega)$ to be zero, we must have $A(\omega) = M(\omega)$ for arbitrary $J(\omega)$. Setting Equation (B-1) equal to Equation (B-2), we obtain

$$\int d\omega''\, J(\omega'')\Big[H_o(\omega)H(\omega'')S(\omega-\omega'')$$
$$-\int d\omega'\, G(\omega,\, \omega')S(\omega'-\omega'')H_o(\omega')\Big] = 0. \qquad (B\text{-}3)$$

For $S(\omega) \simeq e^{-i\frac{\gamma}{2}\omega^2}$ where $\gamma = \dfrac{1}{2\beta}$, Equation (B-3) becomes

$$0 = \int d\omega''\, J(\omega'')e^{-i\frac{\gamma}{2}\omega''^2}\left[H_o(\omega)H(\omega'')e^{-i\frac{\gamma}{2}\left(\omega^2-2\omega\omega''\right)}\right.$$
$$\left. -\int d\omega'\, G(\omega,\, \omega')H_o(\omega')e^{-i\frac{\gamma}{2}\left(\omega'^2-2\omega'\omega''\right)}\right]. \qquad (B\text{-}4)$$

Define $\tilde{G}(\omega,\, \omega') = e^{-i\frac{\gamma}{2}\omega'^2}G(\omega,\, \omega')$ and multiply the square bracket in Equation (B-4) by $\left(d\omega''/2\pi\right)e^{-i\eta\Omega\omega''}$ and integrate to get

$$H_o(\omega)e^{-i\frac{\gamma}{2}\omega^2}\frac{1}{2\pi}\int d\omega''H(\omega'')e^{i\omega''\gamma(\omega-\Omega)} = \frac{1}{\gamma}\tilde{G}(\omega,\, \Omega)H_o(\Omega),$$
$$(B\text{-}5)$$

but $h(t) = \dfrac{1}{2\pi}\int d\omega'\, H(\omega')e^{i\omega' t}$ and if $H_o(\omega) = 1$ for $-B/2 < \omega \le B/2$, then

$$G(\omega,\, \Omega) = \gamma e^{-\frac{\gamma}{2}\left(\omega^2-\Omega^2\right)}h(\gamma(\omega-\Omega)), \qquad (B\text{-}6)$$

where $-B/2 \le \Omega \le B/2$. We have separated out the phase correction terms $S^*(x) = \exp\!\left(-i\gamma x^2\right)$ in the bottom insert in Figure 4.

ANGLE OF ARRIVAL ESTIMATION WITH A POLARIZATION DIVERSE ARRAY

William P. Ballance*, *Ralph A. Coan*

Raytheon Systems Company
Adaptive and Signal Processing Technology Center
2000 East Imperial Highway
El Segundo, CA 90245
William.Ballance@west.boeing.com, rcoan@west.raytheon.com

ABSTRACT

In this paper we derive the Maximum-Likelihood Estimator (MLE) for angle of arrival (AOA) estimation for a general 3-dimensional array with arbitrarily polarized elements. Such an array requires joint estimation of the unknown source polarization as well as its AOA. The key is that polarization enters the model linearly, and hence a closed-form estimate can be computed. The result is that AOA estimation requires a search only over AOA, just as in the known-polarization case. The BWE is shown to be the MLE for a single snapshot, but not for multiple snapshots. The corresponding Cramer-Rao Bound (CRB) is derived and a simple example is used to show agreement. Moreover, an intuitive explanation for the impact that unknown polarization has on the shape of the MLE performance index, and therefore on AOA performance, is provided. Finally, the effect of incorrectly assuming source polarization and attempting "known-polarization" AOA estimation is addressed.

1. INTRODUCTION

This paper investigates the problem of angle of arrival (AOA) estimation of a polarized signal source with a polarization diverse array. This type of array consists of elements with different orientations or polarization responses. The diverse nature of such an array may be the result of physical constraints, or may be purposely introduced to better separate signal sources, providing an additional domain, i.e., polarization, for discrimination. The cost of this additional domain is that AOA estimation requires that the source polarization be jointly estimated.

The paper is organized as follows. Section 2 contains the theoretical analysis of the general problem. The results hold for a signal source of arbitrary polarization in the presence of interference plus noise. The array is also general, with arbitrary element polarization responses and element locations. This includes elements that can be

separated by many wavelengths, as often found in space telescope applications. After presenting the polarization model, the Maximum-Likelihood Estimator (MLE) that jointly estimates AOA and polarization is derived. The MLE for the single-snapshot model is shown to be equivalent to the BWE (Beamform, Whiten, and Energy Detect) technique developed by Steinhardt [1]. The BWE is not the MLE, however, for the multiple-snapshot model. The Cramer-Rao Bound (CRB) on AOA estimation, including polarization, is then derived. Section 3 demonstrates the concepts for an example array configuration. The example is purposely made simple to facilitate an intuitive feel for the mechanics of the MLE and why AOA performance can degrade when polarization must be estimated. The (asymptotic) agreement of the MLE with the CRB is shown. Given that AOA estimation performance can degrade due to the need to estimate polarization, the performance of the known-polarization MLE is addressed when the polarization is incorrectly assumed. Section 4 provides a succinct summary.

2. THEORY

2.1 Problem Formulation

The standard array observation model for a single time-snapshot for an N-element 3-dimensional array is

$$\underline{x} = \underline{d}(\underline{\theta})a + \underline{n}$$

where $\underline{n}$ is complex-Gaussian, zero-mean interference-plus-noise with covariance R. The complex scalar a is the signal waveform value at the array reference point, and the signal direction vector $\underline{d}(\underline{\theta})$ is given by

$$[\underline{d}(\underline{\theta})]_n = f_n(\underline{\theta})e^{j\omega_o\tau_n(\underline{\theta})} \; ; \; n = 1,...,N$$

where $\underline{\theta} = [\phi \; \theta]^T$, ϕ is azimuth, and θ is elevation. The term $f_n(\underline{\theta})$ is the complex-valued element pattern, ω_o is the carrier frequency, and $\tau_n(\underline{\theta})$ is due to element separation.

* Dr. Ballance is now with The Boeing Company.
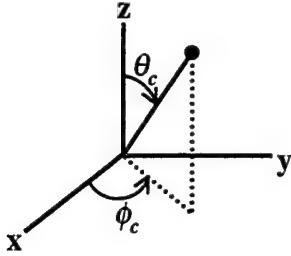
Figure 1. Three-dimensional coordinate system.

Using the coordinate system shown in Figure 1, and denoting the location of the nth element as (x_n, y_n, z_n), the time shift is given by

$$\tau_n(\underline{\theta}) = \frac{1}{c}\left(x_n \sin\phi \cos\theta + y_n \cos\phi \cos\theta + z_n \sin\theta\right)$$

where azimuth $\phi = 90° - \phi_c$ and elevation $\theta = 90° - \theta_c$.

The unknown model parameters are $\underline{\theta}$ and a, for a total of four real-valued unknowns.

The generalization for a polarization diverse array is simply

$$\underline{x} = \underline{d}(\theta, \rho)a + \underline{n}$$

where the polarization state is given by $\underline{\rho} = [\rho_h \; \rho_v]^T$. The direction vector is given by

$$[\underline{d}(\theta, \rho)]_n = f_n(\underline{\theta}, \underline{\rho})e^{j\omega_o\tau_n(\underline{\theta})}.$$

The electric field of a polarized wave traces an ellipse as time progresses. The electric field may be represented as a linear combination of horizontal and vertical polarization

$$\underline{E} = \rho_h\hat{\phi} + \rho_v\hat{\theta}$$

where $\rho_h = |\underline{E}|\cos\gamma$, $\rho_v = |\underline{E}|\sin\gamma\, e^{j\eta}$.

The real-valued quantities γ and η determine the ellipse orientation β and axial ratio $r = (minor\ axis)/(major\ axis)$.

Just as the electric field can be represented in terms of a basis set of polarizations (e.g., horizontal and vertical, RHCP and LHCP, an elliptical polarization and its orthogonal polarization (i.e., "pol" and "cross-pol")), so can the element pattern

$$f_n(\underline{\theta}, \underline{\rho}) = \rho_h f_{h,n}(\underline{\theta}) + \rho_v f_{v,n}(\underline{\theta})$$

as well as the direction vector [2]

$$\underline{d}(\theta, \rho) = \rho_h\underline{d}_h(\underline{\theta}) + \rho_v\underline{d}_v(\underline{\theta})$$

where

$$\left[\underline{d}_h(\underline{\theta})\right]_n = f_{h,n}(\underline{\theta})e^{j\omega_o\tau_n(\underline{\theta})}, \left[\underline{d}_v(\underline{\theta})\right]_n = f_{v,n}(\underline{\theta})e^{j\omega_o\tau_n(\underline{\theta})}.$$

The unknown model parameters are $\underline{\theta}$, γ, η, and a, for a total of 6 real-valued unknowns. The model is linear in polarization when viewed in terms of the parameters ρ_h and ρ_v.

2.2 Maximum-Likelihood Estimator (MLE)

In this section, we derive the MLE using two different methods: an incrementalist approach and from first principles (i.e., from the likelihood function).

2.2.1 MLE Derivation: An Incrementalist Approach

Rather than begin with the likelihood function, we begin with the resultant standard ML performance index and then provide the requisite incremental derivation. Maximizing the likelihood function is equivalent to maximizing the following performance index

$$J(\theta, \rho) = \frac{\left|\underline{w}(\theta, \underline{\rho})^H \underline{x}\right|^2}{\underline{w}(\theta, \underline{\rho})^H R\underline{w}(\theta, \underline{\rho})} \quad \text{where} \quad \underline{w} = R^{-1}\underline{d}(\theta, \rho).$$

The key is that the direction vector is linear in the unknown polarization parameters

$$\underline{d}(\underline{\theta}, \underline{\rho}) = \rho_h\underline{d}_h(\underline{\theta}) + \rho_v\underline{d}_v(\underline{\theta}) = D(\underline{\theta})\underline{\rho}.$$

So that

$$\underline{w}(\underline{\theta}, \underline{\rho}) = \rho_h\underline{w}_h(\underline{\theta}) + \rho_v\underline{w}_v(\underline{\theta}) = W(\underline{\theta})\underline{\rho}$$

where

$$D(\underline{\theta}) = [\ \underline{d}_h(\underline{\theta})\ \ \underline{d}_v(\underline{\theta})\ ],$$

$$W(\underline{\theta}) = [\ \underline{w}_h(\underline{\theta})\ \underline{w}_v(\underline{\theta})\ ] = [\ R^{-1}\underline{d}_h(\underline{\theta})\ \ R^{-1}\underline{d}_v(\underline{\theta})\ ].$$

Define the 2-by-1 vector output of a horizontal and vertical polarization beamformer as $\underline{y}$, and the corresponding 2-by-2 covariance matrix by R_{yy}, so

$$\underline{y}(\underline{\theta}) = W(\underline{\theta})^H \underline{x}, \quad R_{yy}(\underline{\theta}) = W(\underline{\theta})^H RW(\underline{\theta})$$

The performance index may then be written

$$J(\underline{\theta}, \underline{\rho}) = \frac{\left|\underline{w}(\theta, \underline{\rho})^H \underline{x}\right|^2}{\underline{w}(\theta, \underline{\rho})^H R\underline{w}(\theta, \underline{\rho})} = \frac{\left|\underline{\rho}^H \underline{y}(\underline{\theta})\right|^2}{\underline{\rho}^H R_{yy}(\underline{\theta})\underline{\rho}}$$

which is a Rayleigh quotient in the unknown polarization state $\underline{\rho}$. The closed-form ML estimate of polarization is then $\underline{\rho} = R_{yy}^{-1}\underline{y}$. Substituting this into J results in the simple expression

$$J(\underline{\theta}) = \underline{y}(\underline{\theta})^H R_{yy}^{-1}(\underline{\theta})\underline{y}(\underline{\theta}).$$

Hence, the performance index need only be searched over $\underline{\theta}$. Note that J is identical to the BWE [1] performance index and hence the BWE is the MLE. (This assertion was first made by Roland Stoughton of SAIC.)

2.2.2 MLE Derivation: Succinct Proof from First Principles

The model may be written as

$$\underline{x} = \underline{d}(\underline{\theta}, \rho)a + \underline{n} = a_h \underline{d}_h(\underline{\theta}) + a_v \underline{d}_v(\underline{\theta}) + \underline{n}$$
$$= D(\underline{\theta})\underline{a} + \underline{n}$$

where $\underline{a} = [a_h \quad a_v]^T = \rho\, a$.

The key is to realize that this model is analogous to the single-snapshot two-source model

$$\underline{x} = a_1 \underline{d}(\underline{\theta}_1) + a_2 \underline{d}(\underline{\theta}_2) + \underline{n} = D(\underline{\theta}_1, \underline{\theta}_2)\underline{a} + \underline{n}$$

except that $D = [d(\underline{\theta}_1) \quad d(\underline{\theta}_2)]$ and $\underline{a} = [a_1 \quad a_2]^T$ are replaced by $D = [d_h(\underline{\theta}) \quad d_v(\underline{\theta})]$ and $\underline{a} = [a_h \quad a_v]^T$, respectively. The derivation is well-known and readily follows. The likelihood function is given by

$$L = \pi^{-N} \det(R)^{-1} \exp\left\{-(\underline{x} - D(\underline{\theta})\underline{a})^H R^{-1}(\underline{x} - D(\underline{\theta})\underline{a})\right\}.$$

Maximizing the likelihood function is equivalent to minimizing the nonlinear least-squares cost function

$$C = (\underline{x} - D(\underline{\theta})\underline{a})^H R^{-1}(\underline{x} - D(\underline{\theta})\underline{a}) .$$

Whiten for notational convenience

$$C = \left\| \underline{x}_w - D_w(\underline{\theta})\underline{a} \right\|^2 .$$

The problem is a separable, nonlinear least-squares problem. It is nonlinear in $\underline{\theta}$ and linear in $\underline{a}$, allowing a closed-form solution for $\underline{a}$

$$\frac{\partial C}{\partial \underline{a}^H} = -D_w^H(\underline{x}_w - D_w \underline{a}) = 0 \implies \underline{a} = (D_w^H D_w)^{-1} D_w^H \underline{x}_w .$$

Substitution yields

$$C = \left\| (I - P)\underline{x}_w \right\|^2$$

where $P = D_w(D_w^H D_w)^{-1} D_w^H$ is a projection matrix.

Minimizing C is equivalent to maximizing

$$J = \left\| P\underline{x}_w \right\|^2 = \underline{x}_w^H P \underline{x}_w ,$$

or $J = \underline{x}^H R^{-1} D(D^H R^{-1} D)^{-1} D^H R^{-1} \underline{x} = \underline{y}^H R_{yy}^{-1} \underline{y} .$

Q.E.D.

The BWE is the MLE *only* for the deterministic model with a single snapshot. For multiple snapshots, the model is $[\rho_h d_h(\underline{\theta}) + \rho_v d_v(\underline{\theta})]a_k$, which is akin to $\underline{d}(\underline{\theta}_1)a_{1,k} + \underline{d}(\underline{\theta}_2) a_{2,k}$ where $a_{2,k} = \beta\, a_{1,k}$, the fully correlated multipath problem [3] (though for many snapshots, a stochastic-model ML solution may be preferred).

2.2.3 MLE Processing

A block diagram of the MLE for known and unknown polarization is shown in Figure 2. For known polarization, the MLE coherently combines in polarization. For unknown polarization, the MLE effectively noncoherently combines. If desired, the polarization estimate is computed as

$$\underline{\rho} = R_{yy}^{-1} \underline{y}$$
$$\gamma = \tan^{-1}(\rho_v/\rho_h), \quad \eta = \angle(\rho_v/\rho_h).$$

Because the phase must be estimated, it cannot be exploited to achieve the same level of gain as in coherent combination for known polarization.

Note that if adaptive beamforming is followed by coherent processing stages, then two channels per beam ("horizontal" and "vertical") must be carried until they are noncoherently combined after the last stage.
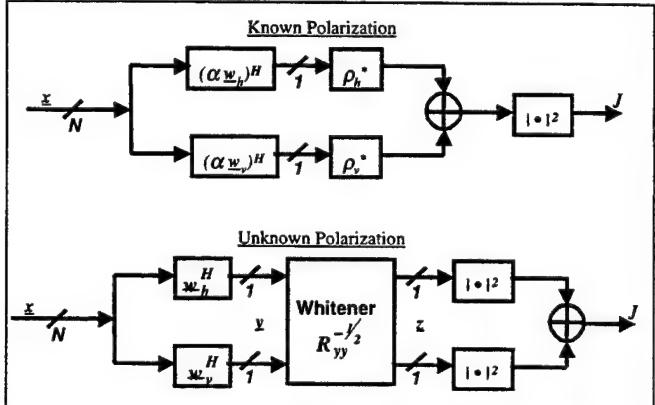
Figure 2. MLE block diagram.

2.3 Cramer-Rao Bound (CRB)

In this section, we derive the CRB on AOA estimation for unknown source polarization. The known-polarization case easily follows. For the single-snapshot observation model $\underline{x} = \underline{h}(\underline{\xi}) + \underline{n}$, where $\underline{\xi}$ is the M-by-1 unknown real-valued parameter vector and $\underline{n}$ is complex-Gaussian with zero mean and known covariance R, the M-by-M Fisher information matrix is given by

$$\Gamma = 2\, Re\,al[H^H R^{-1} H]$$

71

where the mth column of the N-by-M matrix H is $\partial h(\xi)/\partial \xi_m$. For the unknown-polarization problem, we have

$$\underline{h} = \underline{d}(\underline{\theta}, \underline{\rho})a = (\cos\gamma \underline{d}_h(\underline{\theta}) + \sin\gamma\, e^{j\eta} \underline{d}_v(\underline{\theta}))a .$$

The six columns of H are given by

$$\frac{\partial \underline{h}}{\partial \phi} = (\cos\gamma \frac{\partial \underline{d}_h(\underline{\theta})}{\partial \phi} + \sin\gamma\, e^{j\eta} \frac{\partial \underline{d}_v(\underline{\theta})}{\partial \phi})a$$

$$\frac{\partial \underline{h}}{\partial \theta} = (\cos\gamma \frac{\partial \underline{d}_h(\underline{\theta})}{\partial \theta} + \sin\gamma\, e^{j\eta} \frac{\partial \underline{d}_v(\underline{\theta})}{\partial \theta})a$$

$$\frac{\partial \underline{h}}{\partial |a|} = \underline{d}(\underline{\theta}, \underline{\rho})e^{j\angle a}$$

$$\frac{\partial \underline{h}}{\partial \angle a} = j\, \underline{d}(\underline{\theta}, \underline{\rho})a$$

$$\frac{\partial \underline{h}}{\partial \gamma} = (-\sin\gamma \underline{d}_h(\underline{\theta}) + \cos\gamma\, e^{j\eta} \underline{d}_v(\underline{\theta}))a$$

$$\frac{\partial \underline{h}}{\partial \eta} = j\sin\gamma\, e^{j\eta} \underline{d}_v(\underline{\theta})a$$

For the known-polarization problem, only the first four columns are used.

The CRB is given by the elements of $B = \Gamma^{-1}$. Denoting $[B]_{m,n}$ by b_{mn}, the CRB on RMSE for ϕ and θ estimation are given by $\sqrt{b_{11}}$ and $\sqrt{b_{22}}$, respectively. The correlation coefficient between the ϕ and θ estimation is given by $b_{12}/(\sqrt{b_{11}}\sqrt{b_{22}})$. In general, the equi-probability contour (CRB error ellipse) is not aligned with the azimuth or elevation axis. To obtain the CRB for the rotated coordinate system, we compute the eigenvalues and eigenvectors (ordered so that $\lambda_1 \geq \lambda_2$) of

$$M = \begin{bmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{bmatrix}.$$

The quantities $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$ are the CRB on RMSE along the major and minor axes, respectively, while $\beta = tan^{-1}([\underline{u}_1]_2/[\underline{u}_1]_1)$ is the ellipse orientation, where $M\underline{u}_1 = \lambda_1\underline{u}_1$.

Direct numeric computation of the CRB is preferred. While analytic expressions can be obtained using a partitioned-matrix-inverse approach [4], the resultant expressions do not lend great insight due to their complexity. To understand the AOA result, we instead study the shape of the corresponding MLE performance index and its individual components (i.e., whitened horizontal and vertical beamformer responses).
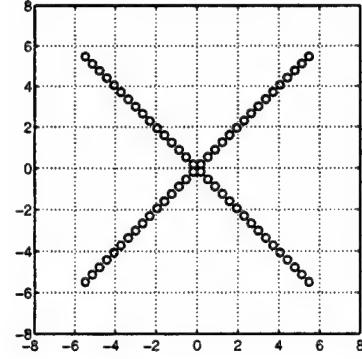

Figure 3. Simple example array configuration.

3. RESULTS FOR EXAMPLE ARRAY

3.1 Array Configuration

A simple example is now considered to provide an intuitive feel for the MLE and illustrate (asymptotic) agreement with the CRB. The array is shown in Figure 3 and consists of two linear segments of dipoles, each with 32 elements and half-wavelength spacing. The dipoles are aligned with the segment axes, which are oriented at $\pm 45°$ with respect to horizontal. The horizontal separation between the phase centers of the segments is variable, but is nominally 14 wavelengths.

3.2 MLE and CRB

Consider a horizontally polarized source at broadside (i.e., 0° azimuth and 0° elevation). To illustrate asymptotic performance (i.e., high SNR), the noise is set to zero so $\underline{x} = (\sqrt{2}/2)[1...1\ 1...1]^T$. The horizontal and vertical direction vectors for broadside are given by

$$\underline{d}_h = (\sqrt{2}/2)[1...1\ 1...1]^T ,$$
$$\underline{d}_v = (\sqrt{2}/2)[1...1\ -1...-1]^T$$

so that at broadside we expect the horizontal beamformer response to peak (since it is matched) and the vertical beamformer response to be zero. The response as a function of azimuth for the beamformers is shown in the bottom part of Figure 4 as dashed lines. As expected, the horizontal beamformer peaks at the source location while the vertical beamformer has a null. Note, however, that as we scan off in azimuth, the vertical beamformer output is nonzero even though the source is horizontally polarized. The upper part of Figure 4 illustrates the corresponding polarization estimate, in terms of the magnitudes of the polarization parameters ρ_h and ρ_v. When $\rho_v = 0$, the estimate is horizontal polarization, but when $\rho_h = 0$, the estimate is vertical polarization. The estimate varies with

72

azimuth in order to maximize the performance index at each look angle. The resultant MLE performance index is shown in the bottom part of Figure 4 as a solid line. The index is the envelope of the horizontal and vertical beamformer responses (actually of the whitened responses, where the whitener accounts for the cross-coupling of the beamformer, though the whitener is simply the identity matrix for this case). The MLE performance index mainlobe is therefore wider than that for known (horizontal) polarization, resulting in larger AOA RMSE. Note that within the mainlobe, the polarization estimate takes on many values. For example, the estimates at the 5 circled points (left to right) are horizontal, RHCP, vertical, LHCP, and horizontal polarization, respectively.

The two-dimensional (azimuth and elevation) MLE performance indices are shown in Figures 5 and 6 for known and unknown polarization, respectively. The equal-height contours, which are elliptical in shape, are shown for each case. The larger ellipse for the unknown-polarization case again indicates that the mainlobe is larger, and therefore the AOA RMSE is larger. The ellipse orientation and axial ratio for each plot agree with the respective CRBs. In addition, Monte Carlo MLE AOA RMSE performance for known and unknown polarization agrees with the respective CRBs. This agreement is expected because the MLE is asymptotically efficient.

Interestingly, the ellipse for the unknown polarization case circumscribes that for the known polarization case. This is true at all angles and source polarizations. The ellipses at 30° azimuth, 30° elevation are shown in Figure 7. While the ellipse size and orientation differ, the unknown case circumscribes that of the known case. This phenomenon is not peculiar to this example, but holds for all arrays.
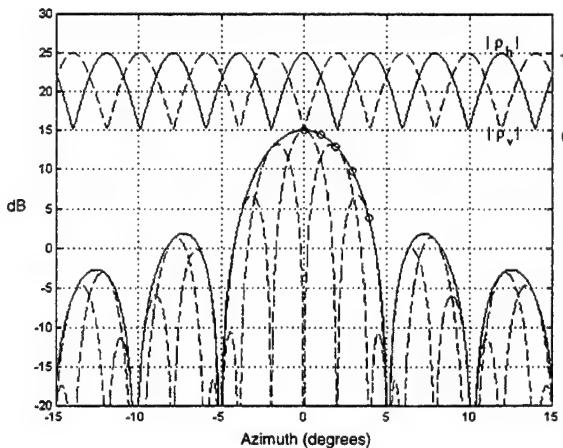

Figure 4. Response of MLE, hzl, and ver beamformers. MLE polarization estimate vs. azimuth is also shown.
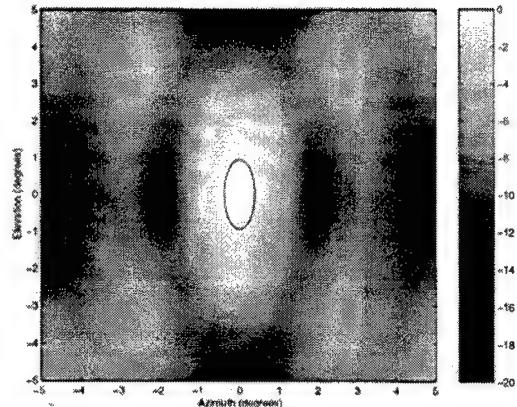
Figure 5. MLE performance index (in dB) for known polarization. Horizontally polarized source at broadside. –0.5 dB contour shown (elliptical).
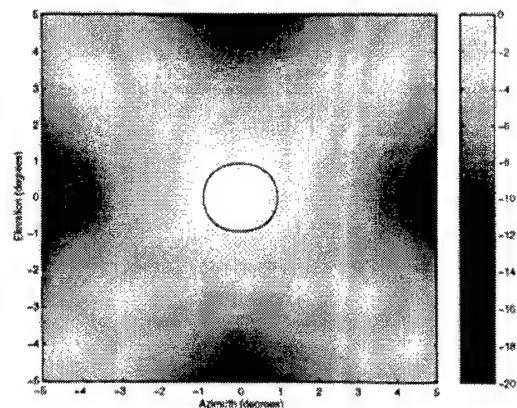

Figure 6. MLE performance index (in dB) for unknown polarization. Horizontally polarized source at broadside. –0.5 dB contour shown (elliptical).
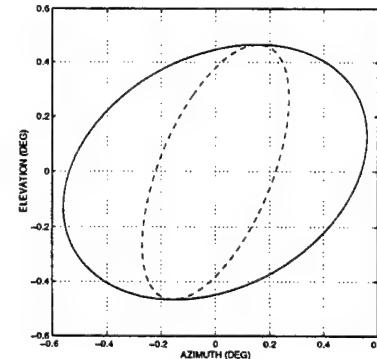

Figure 7. CRB RMSE ellipses for known (--) and unknown (-) polarization. Horizontally polarized source at 30° azimuth, 30° elevation. Absolute plot scaling unimportant.

An interesting observation peculiar to this array is that the AOA RMSE is independent of the separation of the two linear segments when polarization is unknown. When polarization is known, the RMSE decreases with separation because the array aperture is increasing. When polarization is unknown, however, the need to estimate the polarization phase precludes exploitation of the array separation (the performance is the same as known polarization for zero separation). The performance does improve if either array is made to consist of crossed dipoles, or if another linear segment of dipoles, for example, is added.

3.3 Impact of Assumed Polarization

Given that the need to jointly estimate polarization can increase the AOA RMSE, the question arises as to what performance can be achieved using the known-polarization MLE with an assumed polarization. If the assumption is correct, the performance is simply that of the known-polarization case. This section addresses the impact of an incorrect assumption.

The examples considered use the array of Figure 3, with a segment separation of 14 wavelengths, and a source at broadside. In the first case the beamformer assumes the source has circular polarization, but only the horizontal component is truly present. Figure 8 illustrates the resulting performance index. The polarization mismatch between the source and the beamformer is 0.5 dB. More importantly, the azimuth estimate has a 0.82° bias, which is 0.41 of the peak-to-null width of the known (correctly assumed) polarization performance index (0.16 of the unknown). The standard deviation (using Monte Carlo analysis) is commensurate with that of known polarization, but increased by 6% due to the mismatch loss.
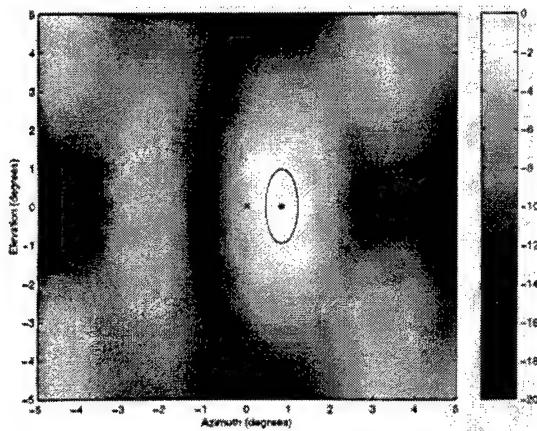

Figure 8. Known-polarization performance index (in dB), assuming circularly polarized source. Horizontally polarized source at broadside. 'x' is true source AOA, '*' is mean of estimated AOA. −0.5 dB contour shown.

In the second case the beamformer assumes horizontal polarization, but in fact the source has a vertical component as well that is 6 dB down with respect to the horizontal component, and 90° out of phase (an elliptical polarization). The resultant mismatch loss is 0.2 dB and the azimuth bias is 0.49°, which is 0.25 of the peak-to-null width of the known-polarization performance index (0.10 of the unknown). Again, the standard deviation is commensurate with that of known polarization. Hence, while the standard deviation can be improved, even a relatively minor error in assumed polarization can produce a significant AOA bias.

4. SUMMARY

We have derived the MLE and CRB for a signal source of arbitrary polarization in the presence of interference and noise, for a general 3-dimensional array of arbitrarily polarized elements. Both the cases of known and unknown polarization were addressed.

Though joint estimation of the source AOA and polarization must be accomplished when the polarization is unknown, the MLE performance index need only be searched over AOA, just as for the known-polarization case. Moreover, we have shown that for the single-snapshot model, the MLE is equivalent to the BWE. This equivalence does not extend to the multiple-snapshot case.

Through the use of a simple example, several important points were made. First, the mechanics of the MLE are intuitively pleasing. Second, AOA performance can degrade due to the need to estimate polarization. Third, the MLE performance index (asymptotically) agrees with the CRB. Fourth, the AOA error ellipse for unknown polarization circumscribes that for known polarization.

Lastly, we showed that the use of the known-polarization MLE with an incorrectly assumed source polarization can produce an AOA error with smaller standard deviation, but with significant bias.

5. REFERENCES

[1] A. Steinhardt, *Proc Adaptive Sensor Array Processing (ASAP) Workshop*, Lexington, MA, March 1994.

[2] E.R. Ferrara, Jr. and T.M. Parks, "Direction finding with an array of antennas having diverse polarizations," *IEEE Trans. Antennas Propagation*, vol. AP-31, no. 2, March 1983, pp. 231-236.

[3] W.P. Ballance and A.G. Jaffer, "Direction finding in the presence of fully correlated specular multipath," *1988 IEEE International Conf. Acoustics, Speech, and Signal Processing*, New York, NY, pp. 2849-2852, April 1988.

[4] W.P. Ballance and A.G. Jaffer, "The explicit, analytic Cramer-Rao bound on angle estimation," *Proc. 22nd Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, pp. 345-351, Nov. 1988.

Distributed Electromagnetic Component Sensor Array Processing

Chong-Meng Samson See *and Arye Nehorai †

Abstract

We propose an approach to localize multiple sources based on spatially distributed electric and magnetic component sensors. By jointly exploiting all available electromagnetic information as well as spatial diversity (time delays), this approach should outperform both a single vector-sensor and scalar-sensor arrays in accuracy of direction of arrival (DOA) estimation.

I. INTRODUCTION

The problem of estimating electromagnetic wave parameters using sensor arrays has attracted significant attention over recent years and lead to the development a number of high resolution algorithms such as MUSIC, ESPRIT and WSF. These algorithms have focussed on direction of arrival estimation in application areas such as wireless communications and radar.

Most existing array processing methods rely on the spatial diversity of the sensor array to estimate the source direction of arrival (DOA). A drawback of this approach is that the accuracy

performance becomes highly dependent on the *electrical* aperture of the array. In many applications, the array is expected to operate over a wide frequency range. To avoid ambiguities in the array manifold, the physical size of such broadband arrays is constrained by the highest operating frequency. Poorer performance at lower frequencies will result due to their larger wavelengths especially when the number of receiver channels is limited. Alleviating this problem by increasing the number of receiver channels to achieve larger "unambiguous" array geometry and size is a costly approach. Another way to overcome this problem is to use multiple sets of sensor arrays where each is optimized to operate over a smaller bandwidth. This may not be feasible in mobile or fast deployment sensor array applications. Hence, there is a need to develop DOA estimation methods that use small aperture array that achieve good performance over a wide operating frequency.

Polarization diversity of the signals can be exploited to improve the DOA estimator performance by estimating the signal polarization parameters along with their DOA using polarization sensitive sensor array [2] [3] [4]. In a recent development, Nehorai and Paldi [1] introduced the concept of vector-sensor array processing where the complete electromagnetic information of the signal is measured and processed. They apply the Poynting relationship between the electric and magnetic measurements to enable estimation of DOA of multiple signal sources with a *single* vector-sensor. Direction finding

with vector sensor (SuperCART antenna array) was recently demonstrated in [7]. As it does not rely on spatial diversity, a DOA estimator using a single vector sensor should exhibit consistent performance over its operating frequency band. When operating as an array of vector sensors, the electromagnetic and time delay measurements can be simultaneously utilized to estimate the DOAs. This allows the use of smaller aperture arrays while maintaining good performance over a wide bandwidth of frequency. However, employing an array of vector sensors is expensive as large number of receivers is needed. For example, an array of 3 vector-sensors array will require a 18-channel receiver.

This paper proposes a simple and effective alternative to achieve good DOA estimation performance with small aperture arrays. The component sensors are distributed as an array of scalar magnetic and electric sensors. It is assumed that the array of scalar magnetic and electric sensors should, in aggregate, measure the all the 3D electric and magnetic components of the electromagnetic wave. We shall call the proposed sensor array as distributed component sensor array. The advantages of this structure are threefold. The full electric and magnetic field components measure by the magnetic and electric sensors can be used to derive the sources' directional information. Secondly, by deploying the magnetic and electric component sensors as a distributed array, additional sources' directional information can be obtained from the differential-delay measurements. Finally, such a sensor array requires a significantly smaller number of receivers to simultaneously utilize the time-delay and *complete* electromagnetic information to perform DOA estimation.

II. MEASUREMENT MODEL

Adopting the conventions in [1], the measurement model of the vector sensor is given by

$$
\begin{bmatrix} \mathbf{y}_E(t) \\ \mathbf{y}_H(t) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_3 \\ (\mathbf{u}\times) \end{bmatrix} \mathbf{V}\mathbf{Q}\mathbf{w}s(t) + \begin{bmatrix} \mathbf{e}_E(t) \\ \mathbf{e}_H(t) \end{bmatrix}, \tag{1}
$$

where

$$
(\mathbf{u}\times) = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_z \\ -u_y & u_x & 0 \end{bmatrix}, \tag{2}
$$

$\mathbf{u}$ is the unit direction vector from sensor to source and u_x, u_y and u_z are the x, y and z components. The matrices $\mathbf{V}$, $\mathbf{Q}$ and vector $\mathbf{w}$ are given by

$$
\mathbf{V} = \begin{bmatrix} -\sin\theta_1 & -cos\theta_1\sin\theta_2 \\ \cos\theta_1 & -\sin\theta_1\sin\theta_2 \\ 0 & \cos\theta_2 \end{bmatrix}, \tag{3}
$$

$$
\mathbf{Q} = \begin{bmatrix} \cos\theta_3 & \sin\theta_3 \\ -\sin\theta_3 & \cos\theta_3 \end{bmatrix} \tag{4}
$$

and

$$
\mathbf{w} = \begin{bmatrix} \cos\theta_4 \\ j\sin\theta_4 \end{bmatrix} \tag{5}
$$

where θ_1, θ_2, θ_3 and θ_4 are the azimuth, elevation, ellipse's orientation and eecentricity angle.

Extending from (1) and assuming that the signal sources are *narrowband*, we can write the measurement model of the distributed component sensor array in multiple sources environment as [6]

$$
\underbrace{\begin{bmatrix} \mathbf{y}_E(t) \\ \mathbf{y}_H(t) \end{bmatrix}}_{\mathbf{y}(t)} = \sum_{k=1}^{d} \mathbf{a}(\boldsymbol{\theta}^{(k)})s_k(t) + \underbrace{\begin{bmatrix} \mathbf{e}_E(t) \\ \mathbf{e}_H(t) \end{bmatrix}}_{\mathbf{n}(t)}, \tag{6}
$$

$$
\mathbf{a}(\boldsymbol{\theta}^{(k)}) = \boldsymbol{\Gamma}(\theta_1^k, \theta_2^k)\boldsymbol{\Omega} \begin{bmatrix} \mathbf{I}_3 \\ (\mathbf{u}_k\times) \end{bmatrix} \mathbf{V}_k\mathbf{Q}_k\mathbf{w}_k
$$

where $\boldsymbol{\theta}^{(k)} = [\theta_1^{(k)}, \theta_2^{(k)}, \theta_3^{(k)}, \theta_4^{(k)}]$ denotes the directional and polarization parameters of the kth source signal. $\boldsymbol{\Gamma}(\theta_1, \theta_2)$ is a diagonal

matrix whose n^{th} diagonal entry is given by $[\mathbf{\Gamma}(\theta_1, \theta_2)]_{nn} = a_n(\theta_1, \theta_2)e^{j\omega_c\tau_n}$ where τ_n is the differential delay of the signal source between the nth component and the phase center and $a_n(\theta_1, \theta_2)$ is the response of the nth component sensor; ω_c is the carrier frequency and $\mathbf{\Omega}$ is a selection matrix of 1 and 0. For example, when orthogonal triads of magnetic and electric sensors are used, $\mathbf{\Omega} = \mathbf{I}_6$. If an additional x-electric component sensor is used, the selection matrix becomes

$$\mathbf{\Omega} = \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ & & \mathbf{I}_6 & & & \end{array} \right]. \qquad (7)$$

From (6), observe that the electromagnetic sources directional information are all embedded in

$$\mathbf{\Gamma}(\theta_1^k, \theta_2^k)\mathbf{\Omega} \left[\begin{array}{c} \mathbf{I}_3 \\ (\mathbf{u}_k \times) \end{array} \right] \mathbf{V}_k.$$

This allows the differential delay measurements due to the spatial diversity of the component sensors and the electromagnetic field measurements to be jointly exploited in estimating the source parameters. Given both the complete electromagnetic and spatial information, good parameter estimation with a smaller aperture array can be expected over a wide frequency range. It suffices to point out that the distributed component sensors array model in (6) generalizes the vector-sensor array [6].

We can express (6) compactly in matrix form as

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \qquad (8)$$

where

$$\mathbf{A} = \left[\mathbf{a}(\boldsymbol{\theta}^{(1)}) \cdots \mathbf{a}(\boldsymbol{\theta}^{(d)}) \right] \qquad (9)$$

and $\mathbf{s}(t) = [s_1(t) \ldots s_d(t)]^T$.

III. CRAMER RAO BOUND

We use the Cramer-Rao bound (CRB) to examine the performance gain achievable by proposed approach. Using the notations, statistical assumptions and results in [1] [6], the CRB is given by

$$\mathbf{C}_{crb}(\boldsymbol{\Theta}) = \frac{\sigma^2}{N} \mathrm{Re}\{\mathbf{J}^{-1}\}, \qquad (10)$$

$$\mathbf{J} = \mathrm{btr}((\mathbf{1} \boxtimes \mathbf{U}) \boxdot (\mathbf{D}^H\mathbf{\Pi}_c\mathbf{D})^{bT})$$

where

$$\begin{aligned} \mathbf{P} &= \mathrm{E}\left(\mathbf{s}(t)\mathbf{s}^H(t)\right), \\ \mathbf{U} &= \mathbf{P}(\mathbf{A}^H\mathbf{A}\mathbf{P} + \sigma^2\mathbf{I})^{-1}\mathbf{A}^H\mathbf{A}\mathbf{P}, \\ \mathbf{\Pi} &= \mathbf{I} - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H, \\ \mathbf{D} &= \left[\mathbf{D}_1^{(1)} \cdots \mathbf{D}_4^{(1)} \cdots \mathbf{D}_1^{(d)} \cdots \mathbf{D}_4^{(d)} \right], \\ \mathbf{D}_l^{(k)} &= \frac{\partial \mathbf{a}(\boldsymbol{\theta}^{(k)})}{\partial \theta_l^{(k)}}, \\ \boldsymbol{\Theta} &= \left[\boldsymbol{\theta}^{(1)T} \cdots \boldsymbol{\theta}^{(d)T} \right]^T \end{aligned}$$

and where σ^2 is the noise power and N is the number of independent snapshots.

IV. NUMERICAL EXAMPLE

We shall illustrate by a numerical example the efficacy of the proposed distributed electromagnetic component sensor array processing as compared with scalar-array processing that uses *electric-only* diversely polarized and co-polarized antenna array. Since the motivation of this development is designing a small aperture sensor array, we shall compare them based on the principle of "equal aperture, equal channels". We assume a six-channel receiver and use a 6 element uniform circular array in this analysis. This will allow the comparison between the performance of a vector-sensor as well as a 6 element diversely and co-polarized array with the proposed distributed EM component sensor array. The diversely polarized array used in this study is an array of x, y and z-electric component sensors. The difference between the diversely polarized and the proposed sensor array is that the former uses only electric component sensors while the latter uses both electric and magnetic component sensors to form a six-element sensor array with

a six-channel receiver. The three sensor arrays are depicted in Figure 1. Note that the inter-element spacing is fixed at $\frac{c}{2f_{\max}}$ where c is the speed of light and $f_{\max}$ is the maximum operating frequency.

An example of the DOA estimation performance as a function of frequency is shown in Figure 2. We considered two uncorrelated sources with $\boldsymbol{\theta}^{(1)} = [1°, 10°, 45°, 0°]^T$ and $\boldsymbol{\theta}^{(2)} = [5°, 9°, -45°, -5°]^T$. The signal to noise ratio is fixed at 10dB. Therein the inter-element spacing of the uniform circular array is fixed at $\frac{c}{2f_{\max}}$. Observe from the figure that the distributed EM sensor array has consistent performance over a wide operating bandwidth. In addition it achieved two orders of magnitude of gain in accuracy of azimuth estimation over the *x-electric* array and one order of magnitude over the *electric only* diversely polarized array at $\frac{f}{f_{\max}} = 0.3$. This result clearly demonstrates the gain obtainable from the full exploitation of the spatial and electromagnetic information by the proposed distributed EM component sensor array.

Figure 3 plots the DOA estimation performance as a function of azimuth angle of separation between uncorrelated two sources of 10dB SNR. The normalized operating frequency is fixed at $\frac{f}{f_{\max}} = 0.3$. The graph shows that the proposed distributed component EM sensor array demonstrates significant performance gain especially for closely spaced sources. This feature is particularly useful in applications with short integration time or at low signal to noise ratio.

V. CONCLUDING REMARKS

We have presented a new approach for the localization of electromagnetic sources by joint exploitation of the spatial diversity and electromagnetic information using spatially distributed electric and magnetic componet sensors. Performance analysis via numerical examples illustrate the potential gain of the proposed approach over scalar and diversely polarized array. The analysis indicates that the proposed distributed component EM sensor array should allow the use of small array apertures while maintaining the desired resolution and accuracy performance over a wide operating bandwidth.

REFERENCES

[1] A. Nehorai and E. Paldi, "Vector-Sensor Array Processing for Electromagnetic Source Localization," IEEE Trans. on Signal Processing, Vol.42, No.2, pp.376-398, Feb. 1994. (a shorter version of this paper has also appeared in Proc 25th Asilomar Conf. Signals Syst. Comput., Nov. 1991, pp. 566-572)

[2] E.R. Ferrara Jr. and T.M. Parks, "Direction finding with an array of antenna having diverse polarization", IEEE Trans. Antennas Propagat., vol. AP-31, pp.231-236, Mar. 1983.

[3] R. Schmidt, "A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation, " P.h.D. dissertation, Stanford University, C.A., Nov. 1981.

[4] I. Ziskind and M. Wax, "Maximum Likelihood Localization of Diversely Polarized Sources by Simulated Annealing," IEEE Trans. Antennas and Propagation, Vol. 38, pp. 1111-1114, July 1990.

[5] J. Li, "Direction and Polarization Estimation Using Arrays with Small Loops and Short Dipoles," IEEE Trans. Antennas and Propagation, Vol. 41, No. 3, pp. 379-387, March 1993.

[6] C.M.S. See and A. Nehorai, "Distributed Electromagnetic Component Sensor Array", full version in preparation

[7] G.F. Hatke, "Performance Analysis of the Super-CART Antenna Array", Project Report AST-22, Lincoln Laboratory, Massachusetts Institute of Technology, 23 March 1992.

Figure 2 – CRB vs Frequency. +: Proposed distributed EM component sensor array. o : Diversely polarized dipole array. -: Vector sensor. ... : Scalar array of omni-directional sensors. × : Scalar array of x-electric sensor (dipole).

Figure 1 – Array Geometry of Distributed EM Component Sensor Array, x-electric (co-polarized) array and electric-only diversely polarized array. $E_x (H_x)$, $E_y (H_y)$ and $E_z (H_z)$ are the electric (magnetic) component sensors.
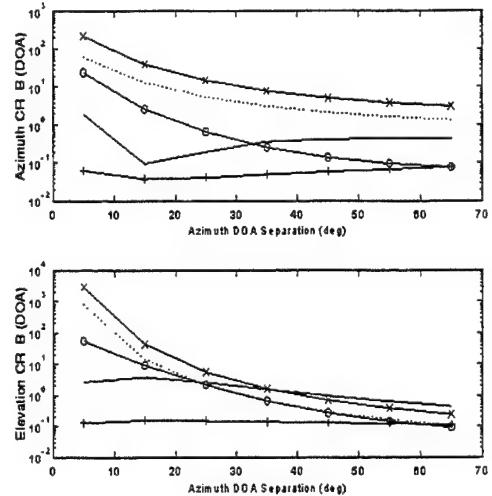

Figure 3 – CRB vs Angular Separation

A COMPUTATIONALLY-EFFICIENT TWO-STEP IMPLEMENTATION OF THE GLRT

Nicholas B. Pulsone, Michael A. Zatman

MIT Lincoln Laboratory
244 Wood Street
Lexington, MA
`pulsone@ll.mit.edu, zatman@ll.mit.edu`

ABSTRACT

In this paper the performance of a new two-step adaptive detection algorithm is analyzed. The two-step GLRT consists of an initial Adaptive Matched Filter (AMF) test followed by a Generalized Likelihood Ratio Test (GLRT). Analytical expressions are provided for the probability of false alarm (P_{FA}) and the probability of detection (P_D) in unknown interference modeled as a multivariate complex Gaussian process. The analysis shows that the two-step GLRT significantly reduces the computational load over the GLRT while maintaining detection and sidelobe rejection performance commensurate with the GLRT.

The two-step GLRT detection algorithm is also compared to another two-step detection algorithm, the Adaptive Sidelobe Blanker (ASB). Both the two-step GLRT and the ASB are characterized in terms of mainbeam detection performance and the rejection of sidelobe targets. We demonstrate that for a given P_{FA} the two-step GLRT has a broad range of threshold pairs (one threshold for the AMF Test and one for the GLRT) that provide GLRT-like mainbeam detection performance. This is in contrast to the ASB, where the threshold pairs that maximize P_D are a function of the target's signal-to-interference-plus-noise ratio. Hence for a fixed pair of thresholds the two-step GLRT can provide slightly better mainbeam detection performance than the ASB in the transition region from low to high detection probabilities.

1. INTRODUCTION

The Adaptive Matched Filter (AMF) [1], Generalized Likelihood Ratio Test (GLRT) [2] and Adaptive Coherence Estimator (ACE) [3] are well-known constant false alarm rate (CFAR) detection schemes. The GLRT is considered to be the benchmark detector for a multivariate complex-Gaussian noise environment. More recently, the Adaptive Sidelobe Blanker (ASB), a two-step test was proposed [4, 5], consisting of an AMF test followed by an ACE test on anything that passes the AMF test. Here we propose a similar two-step implementation of the GLRT, i.e. an AMF test followed by a GLRT. Throughout this paper we refer to the proposed two-step implementation of the GLRT as *2SGLRT* and the one-step GLRT analyzed by Kelly in [2] simply as *GLRT*.

In this paper we show that a two-step implementation of the GLRT achieves significant computational savings with minimal

loss in detector performance. The paper is organized as follows: in Section 2 we describe the various detectors and their computational requirements, followed by an analysis of the 2SGLRT's performance in Section 3, in Section 4 we present simulation results that confirm the validity of our analysis and draw conclusions in Section 5.

2. BACKGROUND

The detection problem addressed in this work is to decide if a test vector x contains a signal vector v where both x and v are $N \times 1$ complex vectors. More specifically, we pose this problem as a decision between two possible hypotheses,

$$x = \begin{cases} n, & \text{noise only hypothesis,} & H_0 \\ n + av, & \text{signal-in-noise hypothesis,} & H_1 \end{cases} \quad (1)$$

where a is an unknown complex scalar. We assume that the $N \times 1$ noise vector n is modeled as a zero-mean complex Gaussian random vector with covariance matrix $\mathbf{R}$ and we assume that $\mathbf{R}$ is unknown to the detector. However, we also assume that the detector has access to a set of noise only data $\{x_i, i = 1, \dots, K\}$ often called training data. It has been shown elsewhere [6] that for the given detection problem (unknown signal multiplier a and unknown covariance matrix $\mathbf{R}$) there is no uniformly most powerful invariant test. Consequently, there are a number of candidate detection algorithms that have been investigated. Three so-called *one-step detectors* which have been proposed are the AMF

$$t_{\text{AMF}} = \frac{|v^H \hat{\mathbf{S}}^{-1} x|^2}{v^H \hat{\mathbf{S}}^{-1} v} \gtrless \eta_{\text{AMF}}, \quad (2)$$

the GLRT

$$t_{\text{GLRT}} = \frac{t_{\text{AMF}}}{1 + x^H \hat{\mathbf{S}}^{-1} x} \gtrless \eta_{\text{GLRT}}, \quad (3)$$

and the ACE

$$t_{\text{ACE}} = \frac{t_{\text{AMF}}}{x^H \hat{\mathbf{S}}^{-1} x} \gtrless \eta_{\text{ACE}} \quad (4)$$

where the superscript ^{H} denotes complex conjugate transpose and $\hat{\mathbf{S}} = \sum_{i=1}^{K} x_i x_i^H$ is the unnormalized sample covariance matrix.

The relationship between the AMF and the other detectors is clear from equations (2), (3) and (4). The GLRT is considered to be the benchmark detector for the multivariate complex-Gaussian signal model. While the ACE detector does not perform as well

at high SINRs [3], for some non-Gaussian noise models the ACE detector is a large dimension approximation of the AMF [7].

Typically the AMF weight vector, $w = \hat{S}^{-1}v$, is computed by performing a QR-decomposition of the data ($8N^2(K - N/3)$ FLOPS) and two back-substitutions ($8N^2$ FLOPS) [8]. Applying the weights to each test vector x under test (a vector inner product) takes $8N$ FLOPS. By comparison the GLRT and ACE require an additional two back-substitutions and inner product for each test vector under test.

The two-step ASB detection algorithm first performs an AMF test on the data followed by an ACE test on the data which passes the AMF test. Thus the additional back-substitutions and vector inner product are only required for a much smaller number of test vectors. Furthermore, the ASB can provide detection performance better than either the AMF or ACE detectors alone by careful choice of thresholds [4]. For a particular signal-to-interference-plus-noise-ratio (SINR), the optimal choice of ASB thresholds provides GLRT-like performance. However, using a fixed pair of ASB thresholds cannot provide GLRT-like performance over all target SINRs.

In a similar fashion, the 2SGLRT proposed here consists of an AMF test on all the data followed by a GLRT test on those test vectors which pass the AMF test (see Figure 1). This vastly reduces the number of back-substitutions and complex multiplies required. Furthermore, we show in the next section that the 2SGLRT has detection performance commensurate with the GLRT proposed by Kelly in [2] for practically all useful scenarios and, in contrast to the ASB, the optimum choice of thresholds is not a strong function of the target SINR.

3. ANALYTICAL PERFORMANCE

In this section we derive expressions for the probability of detection and the probability of false alarm for the 2SGLRT detector. To simplify the analysis we use the following form of the GLRT detector

$$\tilde{t}_{\text{GLRT}} = \frac{1 + x^H \hat{S}^{-1} x}{1 + x^H \hat{S}^{-1} x - t_{\text{AMF}}} \gtrless \tilde{\eta}_{\text{GLRT}}, \qquad (5)$$

where $\tilde{t}_{\text{GLRT}} = (1 - t_{\text{GLRT}})^{-1}$ and $\tilde{\eta}_{\text{GLRT}} = (1 - \eta_{\text{GLRT}})^{-1}$. The test given in (5) is statistically equivalent to the test given in (3).

Our derivation follows the analysis of the ASB given in [4, 5] and uses similar notation. For convenience we relate the distribution of the AMF test statistic t_{AMF} given in (2) and the GLRT statistic $\tilde{t}_{\text{GLRT}}$ given in (5) in terms of two related random quantities, a complex non-central F-distributed random variable $F_{1,L}(\delta_\beta)$ and a complex central β-distributed random variable $\beta_{L+1,N-1}$ as follows

$$t_{\text{AMF}} \overset{d}{=} \frac{F_{1,L}(\delta_\beta)}{\beta_{L+1,N-1}}, \qquad (6)$$

$$\tilde{t}_{\text{GLRT}} \overset{d}{=} F_{1,L}(\delta_\beta) + 1, \qquad (7)$$

where the symbol $\overset{d}{=}$ denotes equality in distribution and δ_β is related to the SINR $\alpha \equiv |a|^2 v^H R^{-1} v$ as $\delta_\beta^2 \equiv \alpha \cdot \beta_{L+1,N-1}$. We assume that v has unit norm, i.e. $v^H v = 1$ and we introduce the integer $L \equiv K - N + 1$. Note that the two random quantities used to describe the the test statistics are related by $\beta_{L+1,N-1}$ which is often referred to as the loss factor [1, 2, 4, 5]. The probability

density function for $\beta_{L+1,N-1}$ is $f_{L+1,N-1}(\beta)$ where $f_{n,m}(\beta)$ is the central beta density defined as

$$f_{n,m}(\beta) \equiv \frac{(n + m - 1)!}{(n - 1)!(m - 1)!} \beta^{n-1} (1 - \beta)^{m-1},$$
$$0 \le \beta \le 1. \quad (8)$$

3.1. Probability of Detection

The 2SGLRT detector declares a target signal present in the test vector x if both the AMF test statistic and the GLRT statistic exceed their respective thresholds, i.e. $t_{\text{AMF}} > \eta_{\text{AMF}}$ and $t_{\text{GLRT}} > \eta_{\text{GLRT}}$. Consequently, the probability of detection (P_D) for the 2SGLRT detector is given by,

$$P_\text{D} = \text{Pr}[t_{\text{AMF}} > \eta_{\text{AMF}}, t_{\text{GLRT}} > \eta_{\text{GLRT}}]. \qquad (9)$$

By using the random quantities given in (6) and (7) we can express the P_D given in (9) conditioned on the loss factor β,

$$P_\text{D} = \int_0^1 \text{Pr}[(F_{1,L}(\delta_\beta) > \beta \eta_{\text{AMF}},$$
$$F_{1,L}(\delta_\beta) > \tilde{\eta}_{\text{GLRT}} - 1)|\beta] \cdot f_{L+1,N-1}(\beta) d\beta, \quad (10)$$

or equivalently

$$P_\text{D} = \int_0^1 \text{Pr}[F_{1,L}(\delta_\beta) > \max(\beta \eta_{\text{AMF}}, \tilde{\eta}_{\text{GLRT}} - 1)|\beta]$$
$$\cdot f_{L+1,N-1}(\beta) d\beta. \quad (11)$$

By defining the parameter γ as follows,

$$\gamma \equiv \begin{cases} 1 & \tilde{\eta}_{\text{GLRT}} - 1 \ge \eta_{\text{AMF}} \\ (\tilde{\eta}_{\text{GLRT}} - 1)/\eta_{\text{AMF}} & \text{otherwise} \end{cases}, \quad (12)$$

we can express the P_D given in (11) as shown

$$P_\text{D} = \int_0^\gamma P_\text{D}^{\text{GLRT}}|\beta \cdot f_{L+1,N-1}(\beta) d\beta +$$
$$\int_\gamma^1 P_\text{D}^{\text{AMF}}|\beta \cdot f_{L+1,N-1}(\beta) d\beta. \quad (13)$$

where $P_\text{D}^{\text{GLRT}}|\beta$ and $P_\text{D}^{\text{AMF}}|\beta$ are the conditional probabilities of detection for the GLRT and the AMF tests respectively. In Figure 2 we illustrate the region of detection for the 2SGLRT that we integrate over in (13). The conditional probabilities given in (13) are expressible as finite sums as noted in [2] and [1]:

$$P_\text{D}^{\text{GLRT}}|\beta = 1 - \frac{1}{(\tilde{\eta}_{\text{GLRT}})^L} \sum_{m=1}^{L} \binom{L}{m} (\tilde{\eta}_{\text{GLRT}} - 1)^m$$
$$G_m\left(\frac{\delta_\beta^2}{\tilde{\eta}_{\text{GLRT}}}\right) \quad (14)$$

$$P_\text{D}^{\text{AMF}}|\beta = 1 - \frac{1}{(1 + \beta \eta_{\text{AMF}})^L} \sum_{m=1}^{L} \binom{L}{m} (\beta \eta_{\text{AMF}})^m$$
$$G_m\left(\frac{\delta_\beta^2}{1 + \beta \eta_{\text{AMF}}}\right). \quad (15)$$

The function $G_i(x)$ in (14) and (15) is defined in terms of the incomplete Gamma function $\Gamma(i, x)$ as follows

$$G_i(x) = \frac{\Gamma(i, x)}{(i-1)!} \equiv e^{-x} \sum_{n=0}^{i-1} \frac{x^n}{n!} \,. \tag{16}$$

Observe that for threshold pairs $(\eta_{\text{AMF}}, \tilde{\eta}_{\text{GLRT}})$ chosen such that

$$\tilde{\eta}_{\text{GLRT}} - 1 \geq \eta_{\text{AMF}} \tag{17}$$

$\gamma = 1$. In this case, the P_D expression given in (13) reduces to the P_D of the one-step GLRT. In other words, with $\gamma = 1$ (or equivalently when $\tilde{\eta}_{\text{GLRT}} - 1 \geq \eta_{\text{AMF}}$) the statistical performance of the 2SGLRT is precisely the same as the one-step GLRT, but with the computational savings outlined in Section 2. On the other hand, with $\gamma = 0$ the 2SGLRT achieves AMF performance and with $0 < \gamma < 1$ the 2SGLRT achieves performance between the GLRT and the AMF.

3.2. Probability of False Alarm

The probability of false alarm (P_{FA}) for the 2SGLRT is obtained from (13) by setting the SINR $\alpha = 0$. Hence, the P_{FA} is as follows

$$P_{\text{FA}} = \int_0^\gamma P_{\text{FA}}^{\text{GLRT}}|\beta \cdot f_{L+1,N-1}(\beta) d\beta +$$
$$\int_\gamma^1 P_{\text{FA}}^{\text{AMF}}|\beta \cdot f_{L+1,N-1}(\beta) d\beta \,, \tag{18}$$

where $P_{\text{FA}}^{\text{GLRT}}|\beta$ and $P_{\text{FA}}^{\text{AMF}}|\beta$ are the conditional probabilities of false alarm for the GLRT and the AMF tests respectively. By substituting $\alpha = 0$ in (14) and (15) and using the Binomial Theorem we obtain the following,

$$P_{\text{FA}}^{\text{GLRT}}|\beta = \frac{1}{(\tilde{\eta}_{\text{GLRT}})^L} \,, \tag{19}$$

$$P_{\text{FA}}^{\text{AMF}}|\beta = \frac{1}{(1 + \beta\eta_{\text{AMF}})^L} \,. \tag{20}$$

Note that the P_{FA} due to the GLRT part of the two-step detection algorithm, i.e. the left hand integral in (18), is independent of the loss factor β and expressible in terms of the incomplete beta function $I_x(p, q)$ [9] as shown

$$\int_0^\gamma P_{\text{FA}}^{\text{GLRT}}|\beta \cdot f_{L+1,N-1}(\beta) d\beta =$$
$$\frac{1}{(\tilde{\eta}_{\text{GLRT}})^L} I_\gamma(L+1, N-1) \,. \tag{21}$$

Furthermore there are an infinite number of threshold pairs $\tilde{\eta}_{\text{GLRT}}$ and η_{AMF} that satisfy (18) for a given P_{FA}.

3.3. Mismatched Signal Detection

In Section 3.1 we derived P_D, the analytical probability of detection for a signal that is aligned with the steering vector, i.e. we assume that the signal vector v given in (1) is aligned with the steering vector v used in (2) and (3). In this section we derive $P_D(\theta)$, the probability of detection for the 2SGLRT to a mismatched signal, i.e. we assume the signal vector v_m arrives from a direction different from the steering vector v. For convenience we define the cosine of the mismatched angle θ between v and v_m in the *whitened space* as follows,

$$\cos^2 \theta = \frac{|v_m^H \mathbf{R}^{-1} v|^2}{(v_m^H \mathbf{R}^{-1} v_m)(v^H \mathbf{R}^{-1} v)} \tag{22}$$

The derivation of $P_D(\theta)$ is similar to the derivation given in Section 3.1. We begin by relating the distributions of the AMF and the GLRT test statistics for a mismatched signal in terms of a complex non-central F-distributed random variable $F_{1,L}(\delta_{\beta(\theta)})$ and a complex non-central β-distributed random variable $\beta_{L+1,N-1}(\theta)$,

$$t_{\text{AMF}}(\theta) \stackrel{d}{=} \frac{F_{1,L}(\delta_{\beta(\theta)})}{\beta_{L+1,N-1}(\theta)} \,, \tag{23}$$

$$\tilde{t}_{\text{GLRT}}(\theta) \stackrel{d}{=} F_{1,L}(\delta_{\beta(\theta)}) + 1 \,, \tag{24}$$

where the non-centrality parameter $\delta_{\beta(\theta)}$ is related to the SINR $\alpha(\theta) \equiv |a|^2 \cdot v_m^H \mathbf{R}^{-1} v_m \cdot \cos^2(\theta)$ as $\delta_{\beta(\theta)}^2 \equiv \alpha(\theta) \cdot \beta_{L+1,N-1}(\theta)$. We assume that v_m has unit norm, i.e. $v_m^H v_m = 1$. Note that the two random quantities used to describe the test statistics are related by $\beta_{L+1,N-1}(\theta)$. The probability density function for the random variable $\beta_{L+1,N-1}(\theta)$ has been derived in [10] as shown

$$f_{L+1,N-1}(\beta; \theta) = e^{-c\beta} \sum_{\ell=0}^{L+1} \binom{L+1}{\ell} \frac{K!}{(K+\ell)!} c^\ell$$
$$f_{L+1,N-1+\ell}(\beta) \,, \tag{25}$$

where $f_{n,m}(\beta)$ is the central beta density defined in (8) and $c = |a|^2 \cdot v_m^H \mathbf{R}^{-1} v_m \cdot \sin^2(\theta)$.

Similar to the derivation in Section 3.1, $P_D(\theta)$ for the 2SGLRT is given by,

$$P_D = \Pr[t_{\text{AMF}}(\theta) > \eta_{\text{AMF}}, \tilde{t}_{\text{GLRT}}(\theta) > \tilde{\eta}_{\text{GLRT}}] \,. \tag{26}$$

Using the parameter γ defined in (12) we can express $P_D(\theta)$ as follows

$$P_D(\theta) = \int_0^\gamma P_D^{\text{GLRT}}(\theta)|\beta(\theta) \cdot f_{L+1,N-1}(\beta; \theta) d\beta(\theta) +$$
$$\int_\gamma^1 P_D^{\text{AMF}}(\theta)|\beta(\theta) \cdot f_{L+1,N-1}(\beta; \theta) d\beta(\theta) \,. \tag{27}$$

where $P_D^{\text{GLRT}}(\theta)|\beta(\theta)$ and $P_D^{\text{AMF}}(\theta)|\beta(\theta)$ are the conditional mismatched probabilities of detection for the GLRT and the AMF tests respectively. In [11] and [1] these conditional mismatched probabilities are expressible as the following finite sums:

$$P_D^{\text{GLRT}}(\theta)|\beta(\theta) = 1 - \frac{1}{(\tilde{\eta}_{\text{GLRT}})^L} \sum_{m=1}^L \binom{L}{m} (\tilde{\eta}_{\text{GLRT}} - 1)^m$$
$$G_m\left(\frac{\delta_{\beta(\theta)}^2}{\tilde{\eta}_{\text{GLRT}}}\right) \tag{28}$$

$$P_D^{\text{AMF}}(\theta)|\beta(\theta) = 1 - \frac{1}{(1 + \beta(\theta)\eta_{\text{AMF}})^L} \sum_{m=1}^L \binom{L}{m}$$
$$(\beta(\theta)\eta_{\text{AMF}})^m G_m\left(\frac{\delta_{\beta(\theta)}^2}{1 + \beta(\theta)\eta_{\text{AMF}}}\right) \,. \tag{29}$$

Note that (27) reduces to (13) when $\cos^2 \theta = 1$, i.e. there is no mismatch between v and v_m.

4. PERFORMANCE EXAMPLES

In this section we show examples of both analytical and simulated performance results for the probability of detection (P_D) with both matched and mismatched signals. These examples assume the following system parameters: the probability of false alarm $P_{FA} = 10^{-6}$, the number of degree of freedom (DOF) $N = 4$ and the number of training vectors $K = 16$.

Figure 3 shows the contours of constant 2SGLRT P_{FA} as a function of the AMF and GLRT thresholds. Figure 4 shows the P_D of the 2SGLRT and ASB detectors as a function of the AMF test threshold for an overall two-stage detector $P_{FA} = 10^{-6}$. For an AMF $P_{FA} \geq 10^{-5.5}, \gamma = 1$ and therefore the threshold pairs of the 2SGLRT satisfy the inequality given in (17), and hence achieve exact GLRT performance for all target SINRs. (Note that the computational savings realized by the ASB and 2SGLRT relative to ACE and the GLRT grow as the AMF test P_{FA} shrinks.) If the AMF test $P_{FA} = 1$ (0 dB - left hand side of the plot), then the 2SGLRT detector becomes the one-step GLRT and the ASB detector becomes the ACE detector. If the AMF $P_{FA} = 10^{-6}$ (-60 dB - right hand side of the plot) then the detector is running as an AMF.

It has previously been reported [4, 5] that for a particular P_{FA} and target SINR there is a single choice of AMF and ACE thresholds which maximizes the ASB's P_D. The fact that the optimal threshold depends on the unknown target SINR α is a drawback of the ASB detector. For the most part the 2SGLRT's P_D is unaffected by the choice of AMF threshold. Only for high AMF thresholds does the 2SGLRT performance begin to degrade. The ASB's P_D quickly degrades from GLRT-like performance when the AMF threshold is either higher or lower than the optimum. As expected, both two-step detectors degrade towards pure AMF performance for high levels of the AMF threshold.

Figure 5 shows analytical and Monte-Carlo simulation (MC) plots of detector P_D as a function of target SINR for the scenario of Figure 4. In Figure 5 the AMF threshold was set to give a P_{FA} of 10^{-5} for the 2SGLRT (thus satisfying (17)) and a P_{FA} of $10^{-5.8}$ for the ASB (thus optimizing the latter's performance for a 23 dB target SINR). Hence in this case the 2SGLRT slightly outperforms the ASB detector at low and medium SINRs. The Monte-Carlo simulation results for the 2SGLRT agree with the theoretical results. In Figure 5 the 2SGLRT is providing the same P_D and P_{FA} performance as the GLRT, but with only the computation of the AMF plus an average of about 1/100000 of the additional computation of the GLRT (assuming that the number of targets is small and snapshots interrogated large).

Figures 6 and 7 show the P_D as a function of cosine squared of the angle ($\cos^2 \theta$) in N dimensional space between the steering vector v in (2, 3 and 4) and the target's steering vector v_m in (1), for the 2SGLRT and ASB detectors respectively. When $\cos^2 \theta = 1$, the target signal is perfectly matched to the steering vector. When $\cos^2 \theta < 1$ some mismatch exists. A small detector P_D for $\cos^2 \theta \ll 1$ shows that the detector is capable of rejecting mismatched targets. This a useful property in many systems (e.g. radar) where strong sidelobe targets may be present.

Initially, as the threshold of the AMF stage in both two-step detectors is reduced, the rejection of mismatched targets increases. As the AMF threshold is further reduced the rejection of mismatched targets by the ASB continues to increase, eventually attaining levels obtained by the one-step ACE detector (comparison with Figure 5 shows that the increased rejection of mismatched

signals also results in a reduced P_D of matched signals - you cannot get something for nothing !). However, the 2SGLRT's rejection of mismatched targets quickly settles to a level commensurate with that of the GLRT as depicted in case C in Figure 6. In fact, the 2SGLRT attains GLRT levels of performance (both matched and mismatched) at approximately the same AMF threshold and AMF P_{FA} as the ASB.

Rather than setting the 2SGLRT thresholds using (17) the results in Figures 3-7 provide a means for setting the 2SGLRT thresholds so that GLRT-like performance is attained with minimal computational expense. Note that the higher the AMF threshold in the 2SGLRT the lower the computational complexity. A radar system example is explained here.

In a radar the initial detection is often followed by some form of non-coherent integration and/or a tracker. There exists some small target SINR so that when the GLRT is used the target's P_D (after any non-coherent integration) would be too low for the tracker to form a track. The AMF threshold of the 2SGLRT is set so that GLRT-like performance is provided by the 2SGLRT for this target SINR, thus ensuring performance commensurate with the GLRT for this and all higher target SINRs. A higher AMF threshold would result in a lower 2SGLRT P_D for the small target, while a lower AMF threshold would increase the required computation.

Figures 8 and 9 give another view of the mismatched signal rejection performance of the 2SGLRT and the ASB respectively. Here a 10-element half-wavelength spaced uniform linear array steered broadside ($\cos \theta_{AZ} = 0$) is modeled. 30 training vectors are available and the P_{FA} for the two-step detectors is set at 10^{-6}. The figures show the P_D for a 20dB SINR target as a function of the cosine of the angle of arrival. The ability of the detectors to reject sidelobe targets is observed. For AMF $P_{FA} > 10^{-5.5}$ the 2SGLRT always gives GLRT-like performance for the 20dB SINR target. (Setting the AMF $P_{FA} > 10^{-4.4}$ implies $\gamma = 1$ and hence results in exact GLRT performance for all target SINRs.) The sideobe rejection of the ASB continues to increase as the AMF threshold falls.

5. SUMMARY

In this paper we have presented a new two-step implementation of the GLRT. The 2SGLRT provides detection performance commensurate with the GLRT while achieving significant computational savings. For the multivariate complex-Gaussian signal model our 2SGLRT has two performance advantages over the other two-step detector examined - the ASB. First, the choice of threshold pairs that maximize P_D for a given P_{FA} does not depend on the unknown target SINR. Second, for a given P_{FA} the detection performance of the 2SGLRT is constant over a wide range of threshold pairs. These two advantages explain the slightly improved detection performance of the 2SGLRT over the ASB in the examples reported in this paper.

It has been demonstrated that altering the ASB's thresholds (while keeping the P_{FA} constant), can increase the ASB's ability to reject 'sidelobe' targets at the expense of mainlobe P_D [5]. This is advantageous in the non-homogeneous scenarios for which the ASB was designed. Assuming the AMF threshold is not set too high, it has been shown here that the 2SGLRT provides sidelobe target rejection commensurate with the GLRT, which is less than the maximum sidelobe rejection afforded by the ASB, but significantly better than that provided by the AMF.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, R. A. Nitzberg, "A CFAR Adaptive Matched Filter Detector," *IEEE Trans. AES*, Vol. 28, No. 1, pp. 208-216 (1992).

[2] E. J. Kelly, "An Adaptive Detection Algorithm," *IEEE Trans. AES*, Vol. 22, No. 1, pp. 115-127 (1986).

[3] L. L. Scharf, L. T. McWholter, "Adaptive Matched Subspace Detectors and Adaptive Coherence," *Proc. 30th Asilomar Conf. an Signals and Systems*, pp. 1114-1117 (1996).

[4] C. D. Richmond, "Statistical Performance Analysis of the Adaptive Sidelobe Blanker Algorithm," *Proc. 31st Asilomar Conf. on Signals and Systems*, pp. 872-876, (1997).

[5] C. D. Richmond, "An Analysis of an Adaptive Detection Algorithm for Non-homogeneous Environments," ICASSP 1998, pp. 2005-2008.

[6] S. Bose, A. O. Steinhardt, "A Maximal Invariant Framework for Adaptive Detection with Structured and Unstructured Covariance Matrices," *IEEE Trans. on Signal Processing*, Vol. 43, No. 9, pp. 2164-2175 (1995).

[7] N. B. Pulsone, "Adaptive Signal Detection in Non-Gaussian Interference," Ph.D. Thesis, *Northeastern University*, May 1997.

[8] G. H. Golub, C. F. Van Loan, *Matrix Computations*, Third Edition, The John Hopkins University Press, Baltimore, MD, (1996).

[9] I. S. Gradshteyn, I. .M Ryzhik, *Table of Integrals, Series, and Products*, Fifth Edition, Academic Press, San Diego, CA (1994).

[10] E. J. Kelly, "Adaptive Detection in Non-stationary Interference, Part III," Technical Report 761, MIT Lincoln Laboratory, Lexington, MA (1987).

[11] E. J. Kelly, "Performance of an Adaptive Detection Algorithm; Rejection of Unwanted Signals," *IEEE Trans. AES*, Vol. 25, No. 2, pp. 122-133 (1989).
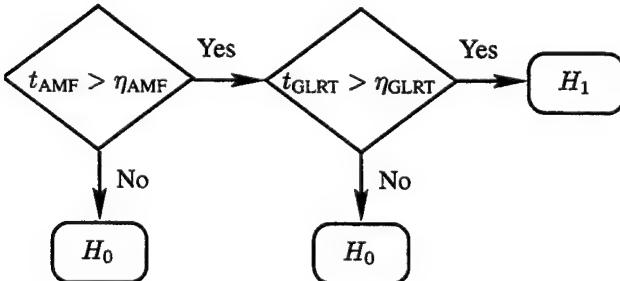


Figure 1: Flow Diagram of the 2SGLRT detection process
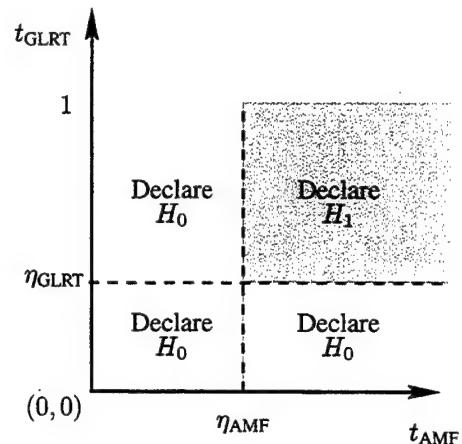
Figure 2: Regions to declare target absent (H_0) and target present (H_1)
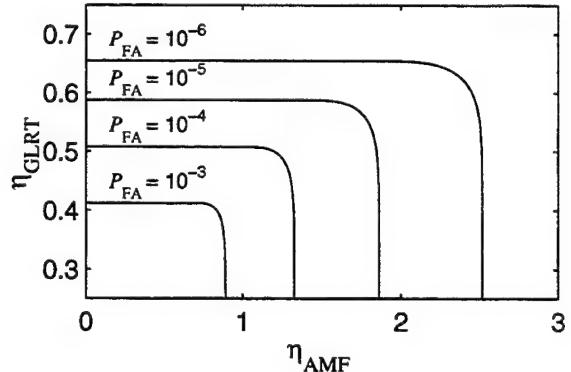
Figure 3: Constant P_{FA} contours for the 2SGLRT detector with $K = 16$ training vectors and $N = 4$ DOFs.
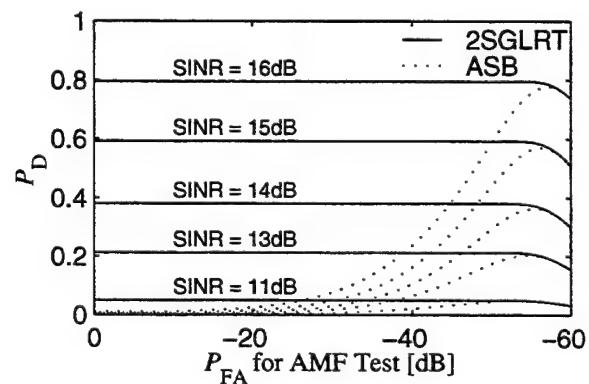

Figure 4: P_D of the 2SGLRT and the ASB detectors as a function of the AMF test threshold for $K = 16$ training vectors, $N = 4$ DOFs and an overall $P_{FA} = 10^{-6}$
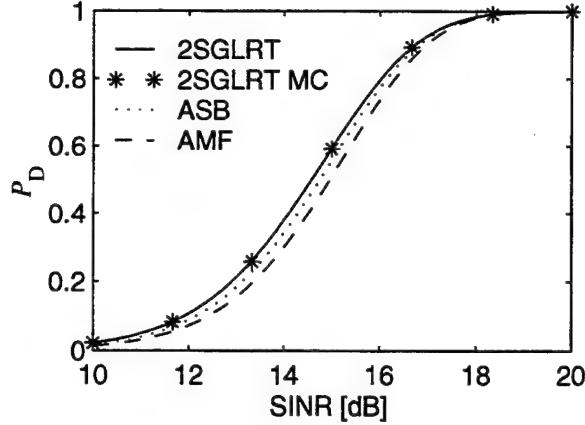
Figure 5: Plots of detector P_D as a function of target SINR for the scenario depicted in Figure 4. AMF P_{FA}s of 10^{-5} for the 2SGLRT and $10^{-5.8}$ for the ASB were used. At the low and medium SINRS the 2SGLRT outperforms the ASB. The '*' represents Monte-Carlo results which confirm the analytical curves.
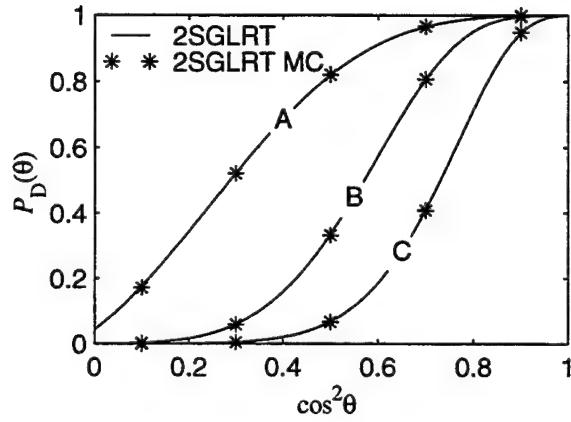

Figure 6: Plots of 2SGLRT $P_D(\theta)$ as a function of $\cos^2\theta$, $P_{FA} = 10^{-6}$, SINR = 20dB, $K = 16$ training vector and $N = 4$ DOFs. P_{FA} for AMF Test (A) -60dB (B) -59.9dB (C) $\geq$ -55dB. The '*' represents Monte-Carlo results which confirm the analytical curves.
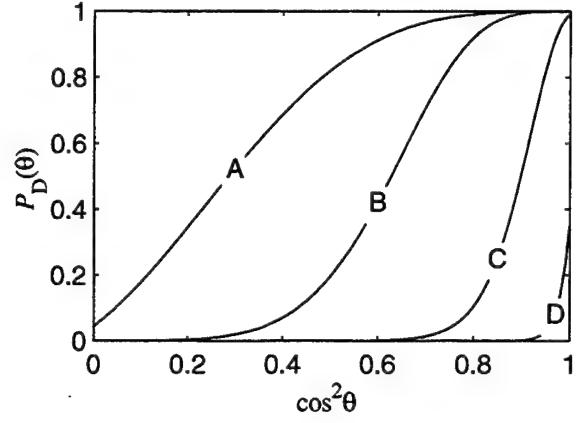

Figure 7: Plots of ASB $P_D(\theta)$ as a function of $\cos^2\theta$, $P_{FA} = 10^{-6}$, SINR = 20dB, $K = 16$ training vectors and $N = 4$ DOFs. P_{FA} for AMF Test (A) -60dB (B) -59.9dB (C) -55dB (D) -30dB.
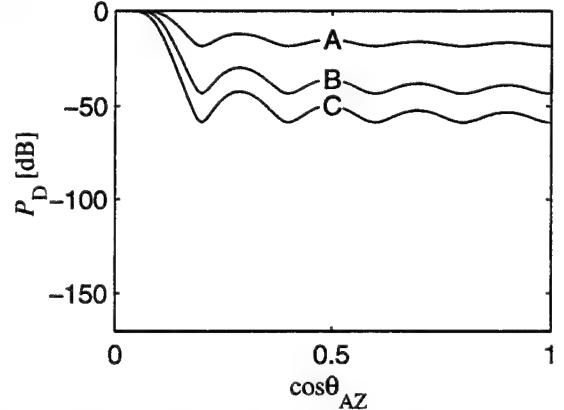
Figure 8: Plots of 2SGLRT P_D as a function of $\cos\theta_{AZ}$, $P_{FA} = 10^{-6}$, SINR = 20dB, $K = 30$ training vectors and $N = 10$ DOFs. P_{FA} for AMF Test (A) -60dB (B) -59.9dB (C) $\geq$-55dB.
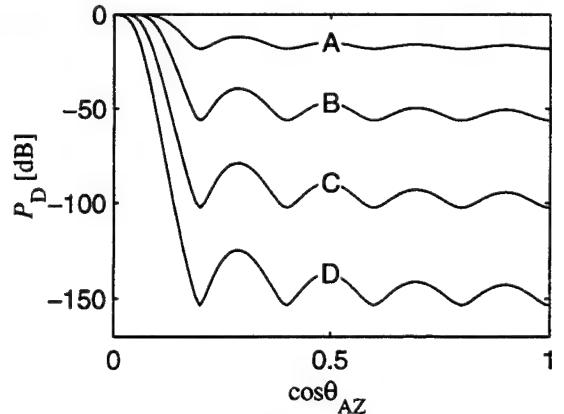

Figure 9: Plots of ASB P_D as a function of $\cos\theta_{AZ}$, $P_{FA} = 10^{-6}$, SINR = 20dB, $K = 30$ and $N = 10$ DOFs. P_{FA} for AMF Test (A) -60dB (B) -59.9dB (C) -55dB (D) -30dB.

THE COSINE GLRT: COMPARISON OF THIS SCALE-INVARIANT GLRT WITH THE KELLY GLRT AND THE AMF

Shawn Kraut and Louis L. Scharf

Dept. of ECE, CB 425
University of Colorado at Boulder
Boulder, CO 80309-0425
`kraut@colorado.edu, scharf@colorado.edu`

ABSTRACT

We examine the problem of "adaptive" detection, wherein the noise covariance structure is unknown, and estimated with training data. We are specifically interested in noise that is not constrained to have the same power level in the test data and training data. For this scenario, we have shown that the "cosine" statistic is the GLRT (Generalized Likelihood Ratio Test) under unknown noise covariance, a companion to the GLRT detector of Kelly. It is invariant to arbitrary scaling of both the training data matrix and the test data. These invariances also make it useful for compound-Gaussian noise scenarios, such as radar clutter modeled by a Gaussian noise process with an additional random amplitude scaling, as proposed by Conte, et. al. We compare the performance of the cosine GLRT, or ACE (Adaptive Cosine Estimator), with the Kelly GLRT and AMF (Adaptive Matched Filter). We have shown that all adaptive detectors of this type have statistically equivalent representations in terms of concise expressions of five statistically independent, scalar random variables, only three of which are needed to completely describe the random fluctuations in the sample covariance. Using these representations, we compare the performance of the adaptive detectors when the noise scaling fluctuates.

1. INTRODUCTION

In this paper we consider the detection of a signal $\underline{\psi}$ embedded in complex Gaussian noise. The detection is based on a multivariate measurement $\underline{y} \sim CN[\mu\underline{\psi}, \sigma^2\mathbf{R}]$. Here μ is the signal gain and σ is the noise scaling, and $\mathbf{R}$ is the noise covariance. In Section 2, we discuss two detectors under the assumption of known covariance $\mathbf{R}$: the matched filter for known noise scaling, and a "cosine" detector for unknown noise scaling. Next we consider the adaptive case, when $\mathbf{R}$ is unknown and estimated from training data. We

discuss three detectors. The first two are the *ad hoc* adaptive matched filter (AMF) and the Kelly detector, which is a true GLRT. The analog to unknown noise scaling in the adaptive case is the presence of an unknown scaling between the training and test data. The GLRT in the presence of such scaling is again a "cosine" detector, often referred to as the adaptive cosine estimator (ACE). We compare the performance of these detection statistics when the the noise scaling is allowed to vary between realizations, effectively generalizing to a compound-Gaussian noise model for the test data. For this performance analysis, we employ "stochastic representations" for the adaptive statistics, equating them to simple expressions involving a small number of independent scalar random variables. We find stochastic representations conditioned on σ. Under high SNR, the Kelly still performs favorably compared to the ACE statistic. However, the Kelly and AMF require knowledge of the statistics of the noise scaling in order to stabilize the probability of false alarm (PFA). The ACE is invariant to scale, and thus the statistics of σ do not affect its false alarm rate.

2. THE NON-ADAPTIVE MATCHED-FILTER AND COSINE DETECTORS

The simplest detection scheme for testing the presence of a target in Gaussian noise is the *matched filter*, wherein the inner product of the the measurement $\underline{y}$ is taken with a signal template $\underline{\psi}$, weighted by the whitening matrix $\mathbf{R}^{-1}$:

$$n = \frac{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{y}}{\sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi}}}. \tag{1}$$

Here $\underline{y} \sim CN[\mu\underline{\psi}, \mathbf{R}]$, with $\mathbf{R}$ known. The real part of this statistic may be compared with a threshold η to decide upon the hypotheses H_0 (signal absent) or H_1 (signal present). The matched filter has a complex normal distribution, $n \sim CN[\mu\sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi}}, 1]$. The detection performance improves

when the mean of this statistic, $\mu\sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1}\underline{\psi}} = \sqrt{SNR}$ increases; thus its square is the effective output signal-to-noise ratio (SNR). When the phase of the signal is unknown, the matched filter may be magnitude squared to yield the matched subspace detector (MSD):

$$\chi^2 = |n|^2 = \frac{|\underline{\psi}^\dagger \mathbf{R}^{-1}\underline{y}|^2}{\underline{\psi}^\dagger \mathbf{R}^{-1}\underline{\psi}}. \tag{2}$$

The MSD is complex chi-squared (or gamma) distributed.

Next we consider the case where we assume knowledge of the noise structure, but not the overall noise level, i.e. $\underline{y} \sim CN[\mu\underline{\psi}, \sigma^2\mathbf{R}]$, with $\mathbf{R}$ known but σ^2 unknown. Then matched filter may be normalized by the magnitude-squared of the measurement, weighted by $\mathbf{R}^{-1}$, to yield

$$cos = \frac{\underline{\psi}^\dagger \mathbf{R}^{-1}\underline{y}}{\sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1}\underline{\psi}}\sqrt{\underline{y}^\dagger \mathbf{R}^{-1}\underline{y}}}. \tag{3}$$

This statistic may be more easily interpreted by considering the whitened measurement and signal vectors $\underline{z} = \mathbf{R}^{-\frac{1}{2}}\underline{y}$, $\underline{\phi} = \mathbf{R}^{-\frac{1}{2}}$; it can then succinctly be expressed as a monotone function of a "complex-t" statistic:

$$cos = \frac{t}{|t|^2 + 1}; \qquad t = \frac{\underline{\phi}^\dagger \underline{\phi}}{\sqrt{\underline{\phi}^\dagger \underline{\phi}}\sqrt{\underline{z}^\dagger \mathbf{P}_{\underline{\phi}}^\perp \underline{z}}}, \tag{4}$$

where $\mathbf{P}_{\underline{\phi}}^\perp = \mathbf{I} - \underline{\phi}(\underline{\phi}^\dagger \underline{\phi})^{-1}\underline{\phi}^\dagger$ is the projection onto the subspace perpendicular to the whitened signal, $\mathbf{I} - \underline{\phi}(\underline{\phi}^\dagger \underline{\phi})^{-1}\underline{\phi}^\dagger$. The t statistic has the same form as n, but it is normalized by a scaled estimate of σ: $\sqrt{N-1}\hat{\sigma} = \gamma = \sqrt{\underline{z}^\dagger \mathbf{P}_{\underline{\phi}}^\perp \underline{z}}$. This makes it have a constant false alarm rate (CFAR) with respect to σ, so we refer to it as the CFAR matched filter, or CFAR MF. This t-form was suggested in the seventies for detection in noise of unknown power level [1, 2, 3].

In the absence of knowledge of signal phase, the "cosine" detector can be magnitude squared to yield the CFAR MSD:

$$\beta = |cos|^2 = \frac{|\underline{\psi}^\dagger \mathbf{R}^{-1}\underline{y}|^2}{(\underline{\psi}^\dagger \mathbf{R}^{-1}\underline{\psi})(\underline{y}^\dagger \mathbf{R}^{-1}\underline{y})} = \frac{F}{F+1}, \tag{5}$$

where $F = |t|^2$ is a scaled complex-F random variable.

The MSD and the CFAR MSD have interesting invariances with respect to transformations of the whitened measurement $\underline{z}$. The MSD is invariant to translations of $\underline{z}$ in the subspace $\langle\underline{\phi}\rangle^\perp$; the CFAR MSD is invariant to rotations in the subspaces $\langle\underline{\phi}\rangle$ and $\langle\underline{\phi}\rangle^\perp$, and to scaling of $\underline{z}$. The MSD and CFAR MSD have been shown to be uniformly most powerful within the class of detectors which share their respective invariances (UMP-invariant) [4]. They are

also Generalized Likelihood Ratio tests (GLRT) [5], obtained by inserting maximum-likelihood (ML) estimates for unknown parameters into the likelihood ratio, given by ratio the probability density function (pdf) of $\underline{y}$ under H_1 to that under H_0. The MSD is obtained by inserting ML estimates for the signal phase and scaling μ, and the CFAR MSD is obtained by inserting ML estimates signal phase and scaling, and also the noise scaling σ.

3. THE ADAPTIVE DETECTORS: AD HOC AND GLRT

In the adaptive scenario, the assumption of known noise covariance $\mathbf{R}$ is relaxed. It is instead assumed that one has access to training data vectors $\{\underline{x}_i\}$ which share the same noise covariance as the test data. A natural way to employ these training vectors is to sum their outer products to form a sample covariance estimate of $\mathbf{R}$, $\mathbf{S} = \frac{1}{K}\sum_i \underline{x}_i\underline{x}_i^\dagger$, where K is the sample support, or number of training vectors. The sample covariance may substituted into the expression for the MSD in Equation 2. This straightforward, though *ad hoc* approach, yields the Adaptive Matched Filter, or AMF, of [6, 7]:

$$\widehat{r^2} = \frac{|\underline{\psi}^\dagger \mathbf{S}^{-1}\underline{y}|^2}{K\underline{\psi}^\dagger \mathbf{S}^{-1}\underline{\psi}}, \tag{6}$$

where we have scaled the denominator by K in order to simplify the statistical distribution of $\widehat{r^2}$.

While the AMF is the natural generalization of the MSD, it is not a generalized likelihood ratio test (GLRT); this is obtained by considering the joint pdf of the training data vectors *and* the test vector, which under H_1 is given by

$$
\begin{aligned}
f_1(\underline{y}, \underline{x}_1, \dots \underline{x}_K) &= f_1(\underline{y})\prod_{i=1}^{K} f(\underline{x}_i) \\
&= \frac{1}{\pi^N \|\mathbf{R}\|} \exp\left\{-(\underline{y} - \mu\underline{\psi})^\dagger \mathbf{R}^{-1}(\underline{y} - \mu\underline{\psi})\right\} \\
&\quad \cdot \prod_{i=1}^{K} \frac{1}{\pi^N \|\mathbf{R}\|} \exp\{-\underline{x}_i^\dagger \mathbf{R}^{-1}\underline{x}_i\}. \tag{7}
\end{aligned}
$$

Inserting maximum-likelihood (ML) estimates of the covariance, $\mathbf{R}$, and the signal gain, μ, yields the GLRT of Kelly [8]:

$$\hat{\lambda}(\underline{y}, \underline{x}_1, \dots \underline{x}_K) = \frac{\max_{\mathbf{R},\mu} f_1}{\max_{\mathbf{R}} f_0} = \left(\frac{1}{1 + \widehat{\chi^2}}\right)^{K+1};$$

$$\widehat{\chi^2} = \frac{|\underline{\psi}^\dagger \mathbf{S}^{-1}\underline{y}|^2}{(\underline{\psi}^\dagger \mathbf{S}^{-1}\underline{\psi})(K + \underline{y}^\dagger \mathbf{S}^{-1}\underline{y})}. \tag{8}$$

We now consider what the noise scaling factor σ^2 means in the adaptive case, when one has access to train-

ing data. A consistent interpretation is that σ is a *relative* scaling of the noise in the the test data, with respect to the training data; that is, $\underline{y} \sim CN[\mu\underline{\psi}, \sigma^2 \mathbf{R}]$, whereas $\underline{x}_i \sim CN[\underline{0}, \mathbf{R}]$. Alternatively, we might say that we are confident about the information the training data gives about the noise *structure* $\mathbf{R}$, but there is still an additional *scaling* σ^2 that is unaccounted for. When the ML estimate of σ^2 is inserted into the likelihood ratio, it yields an adaptive cosine statistic [9]

$$
\hat{\lambda}(\underline{y}, \underline{x}_1, \dots \underline{x}_K) = \frac{\max\limits_{\mathbf{R}, \sigma^2, \mu} f_1}{\max\limits_{\mathbf{R}, \sigma^2} f_0} = \left(\frac{1}{1 + \hat{\beta}} \right)^N ;
$$

$$
\hat{\beta} = \frac{|\underline{\psi}^\dagger \mathbf{S}^{-1} \underline{y}|^2}{(\underline{\psi}^\dagger \mathbf{S}^{-1} \underline{\psi})(\underline{y}^\dagger \mathbf{S}^{-1} \underline{y})} = \frac{\hat{F}}{\hat{F} + 1} \quad (9)
$$

In this case, the true GLRT is in fact the natural generalization of the CFAR MSD of Equation 5, obtaining by simply replacing $\mathbf{R}$ by its sample estimate $\mathbf{S}$. It is often referred to as the ACE detector (for Adaptive Cosine/Coherence Estimator) [10]. It has also been proposed by Conte, et. al. [11, 12], for use in compound-Gaussian noise models, an application we will partially address in the latter part of this paper.

In the remainder of this paper we will compare the performance of these three detection statistics when σ is an unknown parameter that varies for different realizations of $\underline{y}$; thus the measurement is scaled by a fluctuating σ, effectively implementing a compound-Gaussian noise model for the test vector.

4. STATISTICALLY EQUIVALENT STOCHASTIC REPRESENTATIONS

In order to evaluate detection and false alarm probabilities conditioned on σ, we perform a series of matrix transformations, a matrix partitioning, and a change of variables in order to equate these adaptive detection statistics to simple functions of scalar random variables. This method was first presented in detail in [13]; the derivation is only briefly summarized here, for the case of the Kelly statistic. In this derivation, we follow the spirit of Reed, Mallet, and Brennan [14], by considering the complex-Wishart distribution of the sample covariance: $K\mathbf{S} \sim CW[K, N, \mathbf{R}]$, where

$$
f(K\mathbf{S}) = \frac{|K\mathbf{S}|^{K-N}}{J(K\mathbf{S})} e^{-tr(\mathbf{R}^{-1} K\mathbf{S})},
$$

$$
J(K\mathbf{S}) = \pi^{\frac{1}{2} N(N-1)} \Gamma(K) \Gamma(K-1) \Gamma(K-N+1) |K\mathbf{S}|^K. \quad (10)
$$

Applying the whitening transformation as in Equation 4, we write the GLRT as

$$
\widehat{\chi^2} = \frac{|\underline{\phi}^\dagger \mathbf{B}^{-1} \underline{z}|^2}{(\underline{\phi}^\dagger \mathbf{B}^{-1} \underline{\phi})(1 + \underline{z}^\dagger \mathbf{B}^{-1} \underline{z})}, \quad (11)
$$

where $\mathbf{B}$ is a complex-Wishart parameterized by identity, $\mathbf{B} \sim CW[K, N, \mathbf{I}]$. Next we perform a Gram-Schmidt procedure to construct a convenient orthogonal basis from the two vectors of interest, $\underline{\phi}$ and $\underline{z}$. Thus we apply a unitary transformation that sets the first basis vector in the direction of $\underline{\phi}$, and the second in the direction of $\mathbf{P}_\phi^\perp \underline{z}$:

$$
\mathbf{U} = \left[\frac{\underline{\phi}}{\sqrt{\underline{\phi}^\dagger \underline{\phi}}}, \quad \frac{\mathbf{P}_\phi^\perp \underline{z}}{\sqrt{\underline{z}^\dagger \mathbf{P}_\phi^\perp \underline{z}}}, \quad \tilde{\mathbf{U}} \right], \quad (12)
$$

where $\tilde{\mathbf{U}}^\dagger \underline{\phi} = \underline{0}$, and $\tilde{\mathbf{U}}^\dagger \mathbf{P}_\phi^\perp \underline{z} = \underline{0}$. This rotation results in

$$
\widehat{\chi^2} = \frac{|\underline{\delta}^\dagger \mathbf{C}^{-1} \underline{\xi}|^2}{(\underline{\delta}^\dagger \mathbf{C}^{-1} \underline{\delta})(1 + \underline{\xi}^\dagger \mathbf{C}^{-1} \underline{\xi})}, \quad (13)
$$

where $\underline{\delta}^\dagger = [1 \quad 0 \quad \cdots \quad 0]$, and $\underline{\xi}$ has only two non-zero components, which we have already seen: $(\mathbf{U}^\dagger \underline{z})^\dagger = \underline{\xi}^\dagger = [\underline{\xi}_2^\dagger \quad \underline{0}^\dagger]$; $\underline{\xi}_2 = \sigma \cdot [n \quad g]^\dagger$. The first is σn, the matched filter of Equation 1; here n is assumed normalized by σ to give it unity variance, i.e. $n \sim CN[\frac{\mu}{\sigma} \sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi}}, 1]$. The second component, $\sigma g = \sqrt{\underline{z}^\dagger \mathbf{P}_\phi^\perp \underline{z}}$, is the scaled estimate of σ, discussed below Equation 4. Then g^2 has a complex-chi-squared distribution, $g^2 \sim \chi_{N-1}^2[0]$.

Next we partition the inverse of $\mathbf{C}^{-1}$, which is also Wishart; $\mathbf{B} \sim CW[K, N, \mathbf{I}]$:

$$
\mathbf{C} = \left[\begin{array}{c|c} \mathbf{C}_{11} & \mathbf{C}_{21}^\dagger \\ \hline \mathbf{C}_{21} & \mathbf{C}_{22} \end{array} \right]; \qquad \mathbf{C}^{-1} = \left[\begin{array}{c|c} \mathbf{D}^{-1} & * \\ \hline * & * \end{array} \right], \quad (14)
$$

where $\mathbf{D} = \mathbf{C}_{11} - \mathbf{C}_{21}^\dagger \mathbf{C}_{22}^{-1} \mathbf{C}_{21}$, the Schur complement of $\mathbf{C}_{22}$, is Wishart distributed with fewer degrees of freedom, $\mathbf{D} \sim CW[K-N+2, 2, \mathbf{I}_2]$. If we write the elements of $\mathbf{D}$ as

$$
\mathbf{D} = \left[\begin{array}{cc} d_1^2 & d_3^* \\ d_3 & d_2^2 \end{array} \right], \quad (15)
$$

then the Kelly GLRT may be written as

$$
\widehat{\chi^2} = \frac{\widehat{\kappa^2}}{\widehat{\kappa^2} + 1}, \quad \text{where}
$$

$$
\widehat{\kappa^2} = \frac{\left| \frac{n}{g} d_2^2 - d_3^* \right|^2}{(d_1^2 d_2^2 - |d_3|^2) \left(\frac{1}{\sigma^2} \frac{d_2^2}{g^2} + 1 \right)}. \quad (16)
$$

The following transformation produces *statistically independent* random variables, which are complex-chi-squared and complex-normal: $h_1^2 = d_1^2 - \frac{|d_3|^2}{d_2^2} \sim \chi_L^2[0]$, $h_2^2 = d_2^2 \sim \chi_{L+1}^2[0]$, and $h_3 = \frac{-d_3^*}{d_2} \sim CN[0, 1]$. Here $L = K - N + 1$ is an over-training parameter. We then have

$$
\widehat{\kappa^2} = \frac{\left| n + \frac{g}{h_2} h_3 \right|^2}{h_1^2 \left(\frac{1}{\sigma^2} + \frac{g^2}{h_2^2} \right)}. \quad (17)
$$
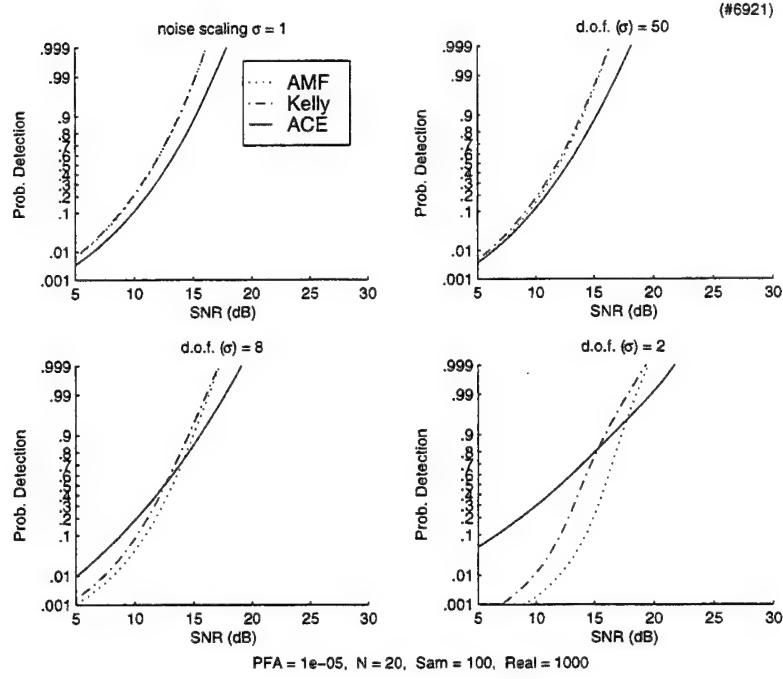
Figure 1: **Detection probabilities for** $K = 100$.

We again associate the parameter label SNR with the mean of n when σ is unity, $\sqrt{SNR} = \mu\sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1}\underline{\psi}}$. By separating out this mean, we have,

$$
\begin{aligned}
\widehat{\kappa^2} &= \frac{\left| m\sqrt{1 + \frac{g^2}{h_2{}^2}} + \frac{\sqrt{SNR}}{\sigma} \right|^2}{h_1{}^2 \left(\frac{1}{\sigma^2} + \frac{g^2}{h_2{}^2} \right)} \\
&= \frac{\left| m + \frac{\sqrt{SNR\cdot b}}{\sigma} \right|^2}{h_1{}^2} \frac{1}{1 + \left(\frac{1}{\sigma^2} - 1 \right) b}, \quad (18)
\end{aligned}
$$

where $m \sim CN[0,1]$ and $b = \frac{h_2{}^2}{h_2{}^2 + g^2}$ is Kelly's beta-distributed loss factor [8, 7]. This expression simplifies considerably when $\sigma = 1$, reducing to the form found in [8].

In a similar manner, the AMF and ACE statistics of Equations 6 and 9 may be written as

$$
\widehat{r^2} = \frac{\left| m + \frac{\sqrt{SNR\cdot b}}{\sigma} \right|^2}{h_1{}^2} \frac{\sigma^2}{b}; \quad (19)
$$

$$
\hat{F} = \frac{\left| m + \frac{\sqrt{SNR\cdot b}}{\sigma} \right|^2}{h_1{}^2} \frac{1}{1 - b}. \quad (20)
$$

5. CONDITIONAL DISTRIBUTIONS

At this point we take a similar approach to that found in [8, 7, 15], and find distribution functions conditioned on the

beta-distributed loss factor, b. The difference here is that we are also conditioning on the noise scaling σ. Conditioned on these parameters, all three statistics are proportional to scaled non-central F statistics, i.e. they are the ratio of a non-central chi-square random variable (in this case, a magnitude-squared complex normal) to another chi-square random variable. We denote the distribution function of the scaled non-central F by $F_{[\nu_n, \nu_d, a]}(\eta)$; this is the probability that the ratio, of a complex-chi-square with ν_n degrees of freedom and a non-centrality parameter a, divided by a complex-chi-square with ν_d degrees of freedom, is less than η. Then the probabilities of detection of our three statistics, conditioned on b and σ, are given by

$$
\begin{aligned}
Pr[\widehat{\chi^2} > \eta | b, \sigma] &= 1 - F_{[1, L, \frac{SNR\cdot b}{\sigma^2}]}\left(\eta\frac{b}{\sigma^2} \right) \\
Pr[\widehat{r^2} > \eta | b, \sigma] &= 1 - F_{[1, L, \frac{SNR\cdot b}{\sigma^2}]}\left(\eta\left\{ 1 + \left(\frac{1}{\sigma^2} - 1 \right) b \right\} \right) \\
Pr[\hat{F} > \eta | b, \sigma] &= 1 - F_{[1, L, \frac{SNR\cdot b}{\sigma^2}]}\left(\eta\left\{ 1 - b \right\} \right). \quad (21)
\end{aligned}
$$

These expressions may be integrated (in practice, numerically) against the pdf's for b and σ to obtain the final probabilities of detection. Kelly has a nice finite-sum expression for the distribution function $F_{[\nu_n, \nu_d, a]}(\eta)$ in [16].

6. NUMERICAL PERFORMANCE COMPARISONS

We compare the detection performance of these three statistics when the noise scaling σ varies between realizations of
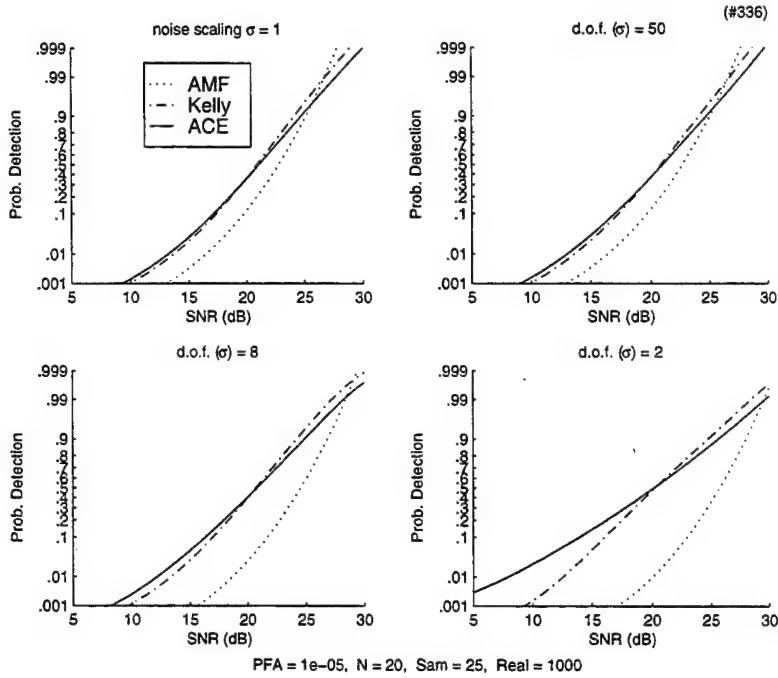
Figure 2: **Detection probabilities for** $K = 25$.

$\underline{y}$. Thus, σ is a random process that scales the multivariate Gaussian process, giving a compound-Gaussian model for $\underline{y}$. Instead of numerical integration, we approximate the integral over the densities of b and σ^2 by averaging realizations of distribution, given by 1000 realizations of b and σ^2 (based on examination of these realizations, we estimate the uncertainty to be about $\frac{1}{100}$ the height of the plots; the qualitative behavior we discuss here is still seen with far fewer realizations). For the density on σ^2, we choose it to be a complex-chi-squared random variable, normalized to have unit mean, and with a variance equal to 1 divided by the number of degrees of freedom.

The resulting detection curves are shown in Figures 1 and 2, for σ^2 constant and chi-squared with 50, 8, and 2 degrees of freedom . They are more easily interpreted by recognizing that the performance of the non-adaptive "cosine" detector of Section 2 only approaches that of the matched filter when N is large compared to the SNR, resulting in a relatively good estimate of the noise scaling, $\sqrt{N-1}\hat{\sigma} = \gamma = \sqrt{\underline{z}^\dagger \mathbf{P}_\phi^\perp \underline{z}}$ (see [4]). Consistent with this observation, we see that the AMF does well against the ACE at high SNRs, and relatively poorly at low SNRs. As the variance of the noise scaling increases, this effect is magnified, and the relative detection performance of the ACE is enhanced further at low SNRs. By comparing Equation 8 with Equations 6 and Equation 9, one can see that the Kelly GLRT approaches the AMF for very high values of the sample support K, and more closely approximates ACE for very

low values of sample support. In Figure 1 the sample support is relatively high, and the detection curve for the Kelly tends to "follow" that of the AMF as the variance of the noise scaling increases; whereas in Figure 2, it more closely follows that of the ACE.

ACE is not only invariant to scaling of the measurement $\underline{y}$, but is a separately invariant to global scaling of the training data set $\{\underline{x}_i\}$ (and thus to scaling of $\mathbf{S}$). For this reason one might expect it to be more robust for small sample support, compared to the AMF. This effect can be seen by comparing the panels of Figure 2 with those of Figure 1. However, for small K, the Kelly is a good approximation to ACE, and still does relatively well.

7. CONCLUSIONS

In the presence of fluctuating noise scaling, the performance of the ACE begins to overtake that of the AMF. The Kelly detector still performs well; while it is not invariant to noise scaling, its normalization term makes it less susceptible to such scaling than the AMF. It should be noted, however, that this comparison has been made by using knowledge of the the distribution of the noise scaling with the AMF and Kelly detectors, to select thresholds for a given probability of false alarm (PFA). The ACE does not require this information; for a set threshold, its PFA does not depend on the statistics of the noise scaling, whereas the PFA of the AMF and Kelly may vary considerably (see Figure 3).

91
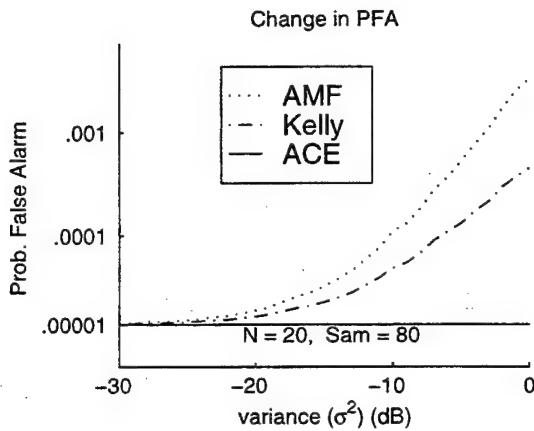
Figure 3: **Change in false alarm probability for a constant threshold**, as the variance of the noise scaling increases.

By introducing a fluctuating noise scaling, we have effectively introduced a compound-Gaussian noise model for the test vector y. This might be considered an intermediate, and analytically tractable, model for the "full" compound-Gaussian scenario considered by Conte, et. al. [11], where the training vectors $\{x_i\}$ also undergo random amplitude scaling, and are thus also compound-Gaussian. In the latter case, the PFA of ACE will no longer be independent of the scaling distribution [11]. But recall that ACE is invariant to global scaling of the whole training data set $\{x_i\}$. For this reason, one might expect the PFA for ACE to be more stable against the scaling distribution in the full compound-Gaussian model; this would be an interesting question to address in future work.

8. REFERENCES

[1] L. L. Scharf and D. W. Lytle, "Signal detection in gaussian noise of unknown level: An invariance application," *IEEE Trans. Information Theory*, vol. 17, pp. 404–411, July 1971.

[2] G. Vezzozi and B. Picinbono, "Detection d'un signal certain dans un bruit spheriquement invariant, structure et characteristiques des recepteurs," *Ann. Telecommun.*, vol. 27, pp. 95–110, 1972.

[3] B. Picinbono, "General detection and estimation theory in an adaptive context," in *Underwater Acoustics and Signal Processing, NATO Advanced Studies Institute Series*, Leif Bjorno, Ed. D. Reidel Publishing, Holland, 1980.

[4] L. L. Scharf, *Statistical Signal Processing*, chapter 4, Addison-Wesley, 1991.

[5] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 2146–2157, 1994.

[6] Wai-Sheou Chen and Irving S. Reed, "A new CFAR detection test for radar," *Digital Signal Processing*, vol. 1, no. 4, pp. 198–214, 1991.

[7] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. A. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 1, pp. 208–216, 1992.

[8] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 1, pp. 115–127, 1986.

[9] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Processing*, accepted for publication.

[10] L. L. Scharf and L. Todd McWhorter, "Adaptive matched subspace detectors and adaptive coherence," in *Proceedings of the 30th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1996.

[11] E. Conte, M. Lops, and G. Ricci, "Asymptotically optimum radar detection in compound-gaussian clutter," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 31, no. 2, pp. 617–625, 1995.

[12] E. Conte, M. Lops, and G. Ricci, "Adaptive matched filter detection in spherically invariant noise," *IEEE Signal Processing Letters*, vol. 3, no. 8, pp. 248–250, 1996.

[13] S. Kraut, L. T. McWhorter, and L. L. Scharf, "A canonical representation for distributions of adaptive matched subspace detectors," in *Proceedings of the 31st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1997.

[14] I. S. Reed, J. D. Mallet, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aero. and Electr. Systems*, vol. 10, no. 6, pp. 853–863, 1974.

[15] Christ D. Richmond, "Statistical performance analysis of the adaptive sidelobe blanker detection algorithm," in *Proceedings of the 31st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1997.

[16] E. J. Kelly, *Finite-Sum Expressions for Signal Detection Probabilities, Technical Report 566*, p. 43, M.I.T. Lincoln Laboratories, 1981.

ADAPTIVE DOPPLER PROCESSING EVALUATED WITH SWAC DATA SETS

James Alsup

SPAWAR Systems Center
53560 Hull Street, Code D712
San Diego, CA 92152
alsup@spawar.navy.mil

Xavier Zabal

Science Applications International Corporation
3990 Old Town Avenue, Suite 303C
San Diego, CA 92110
xavier.a.zabalgogeazcoa@cpmx.saic.com

ABSTRACT

Pulsed single-frequency continuous wave (PCW) sonar echoes in the 1-3 knot range may be difficult to detect in littoral conditions due to the presence of bottom and surface reverberation. The adaptive Doppler processing (ADP) algorithm mitigates this difficulty by using signal-based estimates of the background to suppress reverberation and to help reveal otherwise obscured echoes in this Doppler range. Results are presented using data from the third Shallow Water Active Classification (SWAC-3) sea trial.

1. INTRODUCTION

The Adaptive Doppler Processing (ADP) algorithm has been conceived and developed for the purpose of dealing with sonar echoes obscured by reverberation in littoral scenarios (1-2). Nominally, most of such reverberation is centered about zero Doppler, but has a width and/or offset which is dependent on sea-state and other modulating factors such as current and source or receiver motion. Doppler-sensitive signals, such as the single-frequency pulsed continuous wave (PCW) signal, have been used to allow processors to reject such reverberation while retaining echoes which result from non-zero Doppler targets. ADP is designed to perform this function even more effectively.

The ADP algorithm provides benefit when (a) the environment is dominated by bottom reverberation, and (b) the target echo has a Doppler shift which is large enough to be "non-zero" yet small enough to be obscured by the "zero-Doppler ridge" (this may occur, for example, in the range of 0.5-5knots). Modeling studies indicate that the ADP usefulness declines with increasing center frequency, and we have noted positive performance gains at frequencies up to 1kHz.

2. CONVENTIONAL DOPPLER PROCESSING

The conventional Doppler processor can be written as a filter vector, s_v, operating on a block data matrix, D_p:

$$G_C(v,p) = s_v^* D_p$$

where v is the Doppler index, and p is the order of the data matrix. The filter vector is the sampled Doppler-distorted signal replica:

$$s_v = \left[e^{i\omega_o t(1+2V/c)} \right], \quad t \in [0,...,(N-1)\Delta t]$$

The data matrix appears as:

$$D_p = \begin{pmatrix} x_1 & \cdots & x_{1+p} \\ \vdots & \ddots & \vdots \\ x_N & \cdots & x_{N+p} \end{pmatrix}$$

Note that each new column in D_p contains one new sample and does not contain one old sample in comparison with the previous column. For PCW transmissions, the first equation can be performed via a fast Fourier transform.

As an example, let us consider a 720Hz PCW signal of duration 4 seconds whose (downshifted and decimated) sample rate is 17.75Hz. Then, we could choose N = 71 so that one column in D_p could contain the entire replica. A nominal choice (see below) is p = 4, resulting in 5 columns. The index v might range from -40 through +40, a total of 81 values, which would accommodate a ±10-knot Doppler range with a resolution of 0.25-knot. A choice is typically made to average the values of G_c over these (p+1) samples, resulting in a single output vector of size 81x1 which is then updated at a rate of 3.55Hz, once for each new data matrix.

3. ADAPTIVE DOPPLER PROCESSING

3.1 Minimum variance version

The ADP algorithm is a modification of the minimum variance distortionless response (MVDR) spectrum processor. For the MVDR processor the filter vector $\mathbf{w}_v$ minimizes the total output power subject to the constraint that the data components similar to the desired signal $\mathbf{s}_v$ are not degraded by this filter:

$$\mathbf{G}_{MVDR}(v,p) = \mathbf{w}_v^* \mathbf{D}_p$$

and

$$\left|\mathbf{G}_{MVDR}(v,p)\right|^2 = \mathbf{w}_v^* \mathbf{R}_p \mathbf{w}_v$$

where $\mathbf{R}_p = \mathbf{D}_p \mathbf{D}_p^*$ (the covariance) and $\mathbf{w}_v^* \mathbf{s}_v = 1$ (the constraint).

The determination of this filter vector, $\mathbf{w}_v$, normally requires that the order of the data matrix must exceed the length of the transmitted signal (typically $p > 2N$). However, when the data is dominated by reverberation, it may not be stationary over the interval, and the MVDR solution may become meaningless.

3.2 Reduced rank version

This leads us to the use of rank-deficient methods for determining $\mathbf{w}_v$, and the use of smaller values (say $4 < p < 15$) for the order of the data matrix. A modified processor may be obtained by augmenting the sample covariance matrix using the diagonal loading technique. This permits solving for a new filter vector $\mathbf{u}_v$ directly via singular value decomposition (SVD) of the data matrix. Using this method, only $p+1$ eigenvalues and eigenvectors are determined:

$$\mathbf{G}_{ADP}(v,p) = \mathbf{u}_v^* \mathbf{D}_p$$

and

$$\left|\mathbf{G}_{ADP}(v,p)\right|^2 = \mathbf{u}_v^* \widetilde{\mathbf{R}}_p \mathbf{u}_v$$

where $\widetilde{\mathbf{R}}_p = \mathbf{D}_p \mathbf{D}_p^* + \varepsilon \mathbf{I}$ (the covariance) and $\mathbf{u}_v^* \mathbf{s}_v = 1$ (the constraint), and ε is the white noise factor which determines the strength of the constant term added to the diagonal of the otherwise-rank-deficient covariance matrix. This permits that matrix to be inverted when otherwise it would be unstable due to insufficient rank ($p \ll N$). Trials with experimental data have shown $4 < p < 15$ to be a reasonable range beyond which

reverberation behaves more like noise than like an echo at zero-Doppler. The filter vector is given by:

$$\mathbf{u}_v = \frac{\widetilde{\mathbf{R}}_p^{-1} \mathbf{s}_v}{\mathbf{s}_v^* \widetilde{\mathbf{R}}_p^{-1} \mathbf{s}_v}$$

The inverse diagonally-weighted covariance matrix can be computed as [2]:

$$\widetilde{\mathbf{R}}_p^{-1} = \varepsilon^{-1}\left(\mathbf{I} - \sum_{i=1}^{p+1} \frac{\lambda_i}{\lambda_i + \varepsilon} \mathbf{v}_i \mathbf{v}_i^*\right)$$

where $\mathbf{v}_i$ are the eigenvectors, and λ_i are the corresponding eigenvalues. The first $p+1$ eigenvalues and eigenvectors for each Doppler are computed directly from the data matrix using the SVD algorithm, then used explicitly in the computation of $\widetilde{\mathbf{R}}_p^{-1}$ and $\mathbf{u}_v$. This algorithm is much more efficient than a previous version [3], since the matrix inversions can all be computed via the SVD rather than through cumbersome multi-matrix manipulations.

3.3 Reduced rank/desensitized version

One additional algorithm refinement is required to deal with signal mismatch. This situation occurs when the Doppler replica is not exactly matched to the echo or reverberation being analyzed, because of the finite number of replicas used and the resultant gap between adjacent Doppler bins. In this case some or all of the echo or reverberation signal structure in a given data matrix may be adaptively nulled when in fact it is "closer" to that replica than any other. Rather, what is desired is to null out echo or reverberation components which lie outside the Doppler regime spanned by the separation between successive Doppler replicas. This problem can be rectified by a method known as signal desensitization [4], which allows the algorithm to preserve echo or reverberation energy which is "close to" the frequency of the Doppler replica in use. This is done by defining a modified eigenvalue as follows:

$$\widetilde{\lambda}_i = \lambda_i\left(1 - \left|\mathbf{v}_i^* \mathbf{s}_v\right|^2\right)$$

Then,

$$\widetilde{\mathbf{R}}_p^{-1} = \varepsilon^{-1}\left(\mathbf{I} - \sum_{i=1}^{p+1} \frac{\widetilde{\lambda}_i}{\widetilde{\lambda}_i + \varepsilon} \mathbf{v}_i \mathbf{v}_i^*\right)$$
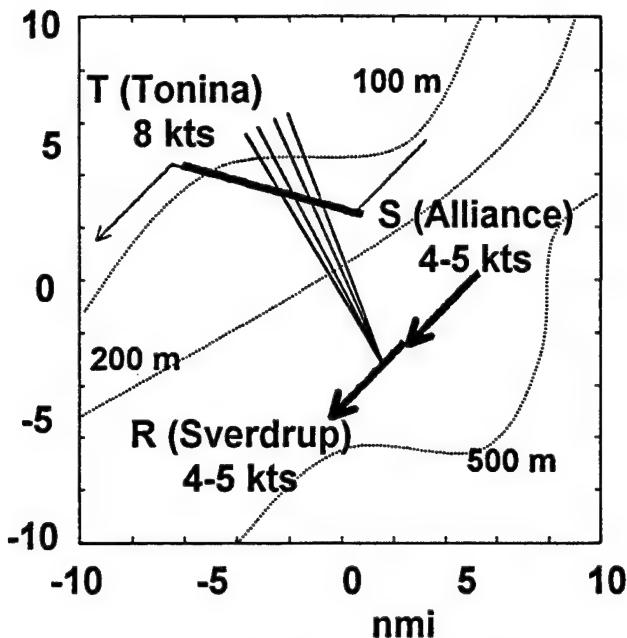
Figure 1. SWAC-3 venue, Run Bravo, near Valencia, Spain: towed source (Alliance) and towed-array receiver (Sverdrup) both heading SW at 4.5kt, offset on same track. Target (Tonina) proceeding SW, WNW, and SW on three 1-hour legs at 8 knots. Water depth varied from 100m to 500m, sloping down to SE.

This action can be seen to be a reduction in the size of a given eigenvalue if the corresponding eigenvector lies near the replica vector used for a given Doppler computation. The smaller modified eigenvalue in turn reduces the contribution of its corresponding eigenvector to the summation in the calculation of the inverse covariance matrix, thus diminishing the amount of nulling applied for echo or reverberation components near the replica's frequency.

When the ADP algorithm with desensitization is applied to actual sonar data, it is observed that both the echo level and the reverberation structure of the "zero-Doppler" ridge are preserved, whereas without this modification the algorithm tends to suppress these real structures which are otherwise "too close" to the replica to avoid being nulled out.

4. ADP EXPERIMENTAL RESULTS

4.1 SWAC-3 venue

Data collected during the Shallow Water Active Classification test #3 (SWAC-3) experiment was used to evaluate different variants of the ADP algorithm. During Run Bravo of this exercise the towed source, towed
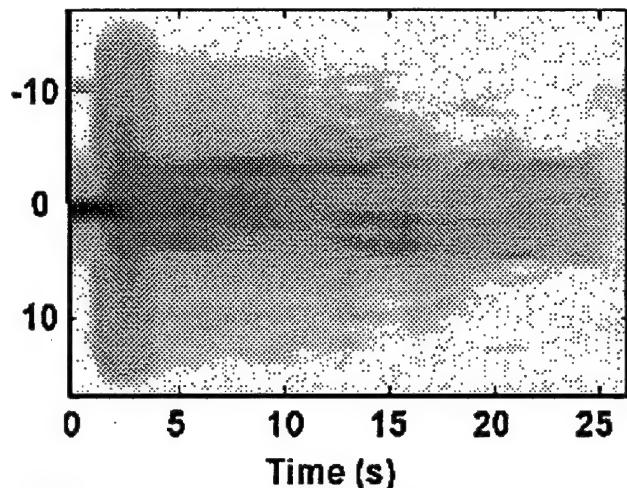

Figure 2. Averaged phone spectrogram from Sverdrup towed array, shows ±18 knots of Doppler on vertical axis versus time in seconds on horizontal axis. Intensities cover a 50dB grayscale range, loudest at the direct blast, then diminishing with both time and Doppler.
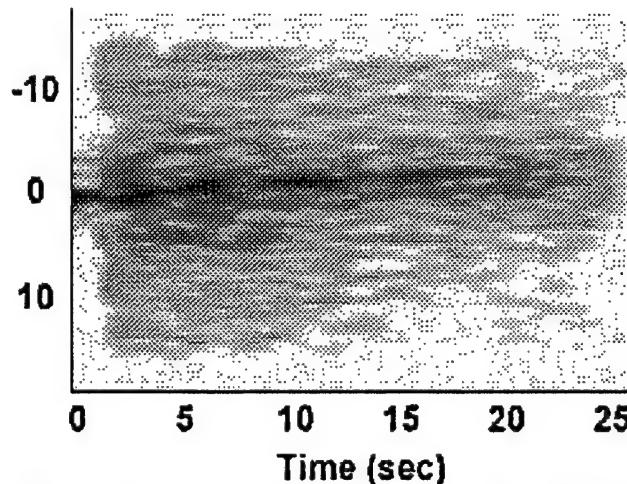

Figure 3. Beam spectrogram from Sverdrup towed array, shows ±18 knots of Doppler on vertical axis versus time in seconds on horizontal axis. Intensities cover a 50dB grayscale range, loudest at the start, then diminishing with both time and Doppler.

receive array, and target followed tracks and ran at speeds and headings as shown in Figure 1. The total duration of the run was 3 hours; each leg of the target track took 1 hour. The Towed Vertically Directive Sonar (TVDS) source transmitted a 4-second CW pulse, centered at 720 Hz, and at intervals of 60 seconds. The subset of the receive array analyzed here consisted of 48 phones spaced at about .45 λ.
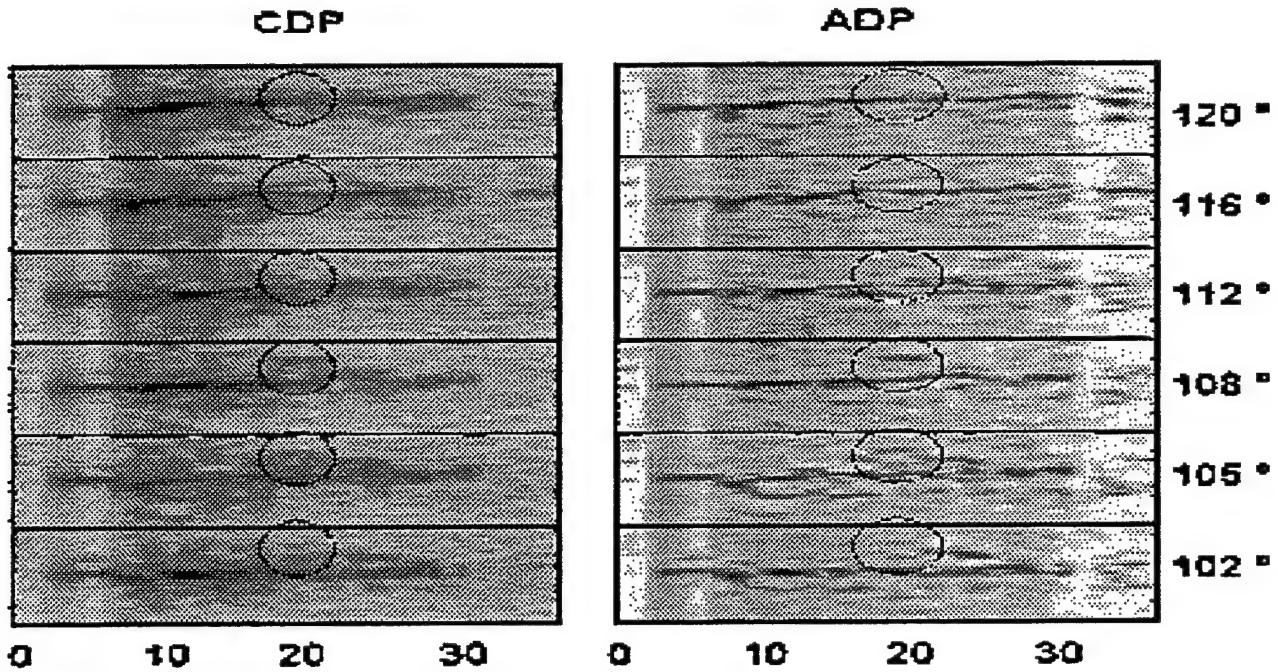
Figure 4. CDP and ADP beam spectrograms from a single ping, showing ±10 knots of Doppler on each of 6 stacked vertical axes, versus time in seconds on the horizontal axis. Intensities cover a 50dB range, from low (white) to high (black).

The data presented here corresponds to the middle leg of the track, covering the period from 1400 to 1500 hours, where the target had a Doppler corresponding to an opening speed of about 6 knots. A total of 4 beams (indicated in the figure) contained the target at all times, though at least six beams were usually processed.

4.2 Phone average spectrogram

A spectrogram averaged over all 48 hydrophones is shown in Figure 2. For this particular ping the data shows some isolated strong echoes at different opening/closing speeds and a continuum of reverberation spread over the ±4.5 knot region as expected for a source and receiver moving at this speed.

4.3 Beam spectrogram

An ideal beam isolates the reverberation coming from a narrow strip on the bottom and therefore acts as a Doppler filter. The spectrogram presented in Figure 3 shows that, indeed, the reverberation is concentrated along a narrow ridge with strong off-Doppler contributions due to leakage through the beam sidelobes. The beam energy appears to have a Doppler spread which is larger than what would be expected. This spread is probably due to array deformation and could have an impact in Doppler detection at low opening/closing speeds. From this figure

it can be seen that the reverberation ridge gradually decreases in Doppler with increasing time delay. This effect is expected for a bistatic sonar configuration, and must be compensated when trying to remove own-ship-Doppler from the display. Only for beams in the direction of tow (fore and aft) can a constant bottom-reverberation Doppler offset versus time be observed.

4.4 CDP and ADP Spectrograms

Figure 4 illustrates a typical result from processing a single ping of SWAC-3 data using both CDP and ADP methods. Six beams are shown. The vertical axis represents Doppler shift in knots, from −10 to +10. The horizontal axis represents time-delay after the direct blast, in seconds. The black ellipse in each spectrogram marks the spot where expected speed (-6kt) and delay (18 sec) of the target echo is supposed to be seen. This echo peaks up in the 4th beam (108 degrees), and can be seen with much greater clarity in the ADP spectrogram. Aggregate measurements of signal excess from several pings are presented below.

4.5 ADP Processing Chain

One of the problems encountered in Doppler detection is the additional dimension in search space that must be
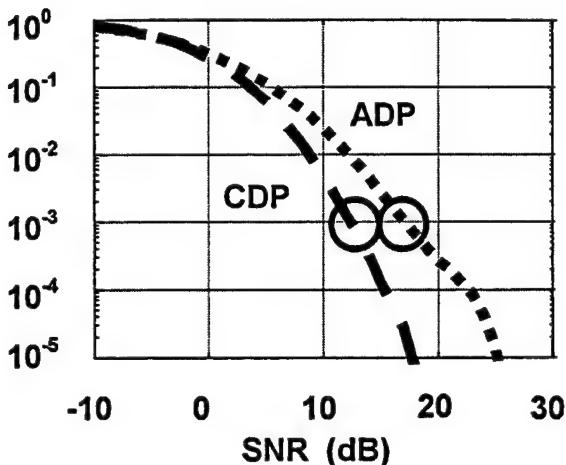
Figure 5. CFAR Thresholds: Lower curve: CDP, upper curve: ADP. Based on PCW data from SWAC-3 bistatic towed array, shows probability of false alarm P_{FA} (vertical axis) versus SNR threshold for Doppler regime between +5 and +10 knots.

presented to the operator. One method to reduce this dimensionality is to combine beams and Doppler regions, as follows: (a) after demodulation, decimation, and CBF are done, each beam is fed to an adaptive Doppler processor module, whose output is a set of Doppler channels; (b) for each Doppler channel a set of beams are combined in an OR operation (in the results presented later we have OR'd over 4 beams covering an angular sector of 14 degrees); (c) the OR'd outputs are normalized by scaling each point in delay and Doppler by a local background average (this output is a measure of SNR); (d) the normalized time delay/Doppler surface is OR'd over a Doppler interval, $[v_1, v_2]$, to provide a single detection function for this region (in the results presented later we have OR'd over a Doppler interval of [-10 –5] knots); when the proper threshold is applied to the beam- and Doppler-OR'd data the output becomes a measure of signal excess, SE; (e) the SE scans are combined across pings to present a 3-dimensional time-history of the Doppler-processed results.

4.6 CFAR Thresholds

In order to compare the performance of different processors, we measured the distribution of this output in the absence of a target to determine a threshold for constant false alarm operation (CFAR), because the probability of false alarm (P_{FA}) distribution depended on whether the conventional or adaptive Doppler processor was used. A total of 110 pings were analyzed to determine this distribution, from all three legs of Run Bravo. Four adjacent beams were Doppler processed,
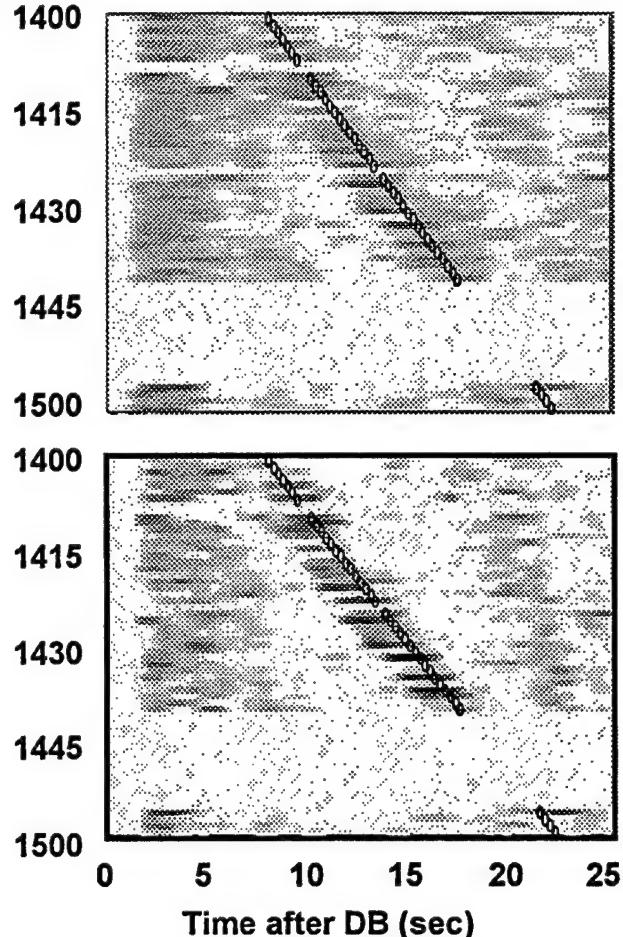

Figure 6. SE Performance: Upper: CDP, DT=12dB; Lower: ADP, DT=16dB. Shows 60 pings (vertical axis) over one-hour period from SWAC-3 bistatic towed array data, 4-second PCW signal. Pings 41-56 are blank. Shows time delay in seconds (horizontal axis) from start of direct blast. Shows signal excess in dB over a [-10 to +20dB] grayscale range.

OR'd, and normalized using a two-dimensional delay-and-Doppler filter. The SNR output was OR'd over the [5,10] knot Doppler range where there were no targets (during the whole run the target was either at 0 knots or close to –6 knots). The P_{FA} distribution of these output data is shown in Figure 5 for both the CDP and ADP processors. At a P_{FA} level of 10^{-3}, the detection threshold associated with CDP is 12dB, and that with ADP is 16dB.

4.7 Track Performance

Using these detection thresholds and the strategy of ORing in both beam and Doppler, we can display the output SE for a set of pings between 1400 and 1500 hours corresponding to the middle leg of the target track. The
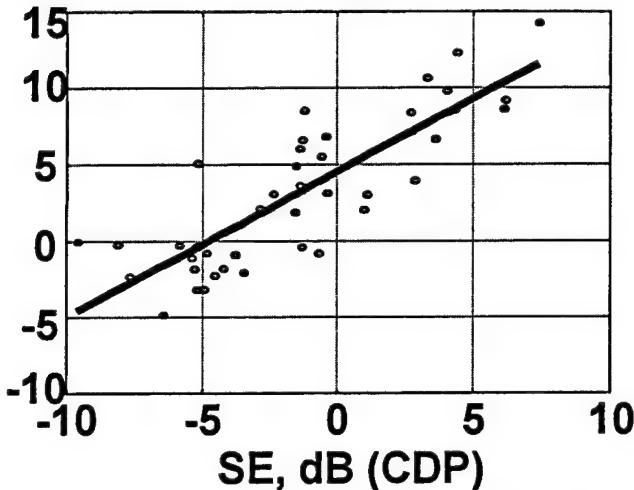
Figure 7. Gain Performance: ADP Gain = SE_{ADP} - SE_{CDP}, is plotted versus SE_{CDP}. SE's are measured from 41 PCW pings from SWAC-3 bistatic towed array data (circle symbols).

position of the target, as expected from the track plan, is indicated in Figure 6 by the small black circles. From this display it can be seen that the ADP target track is much stronger, although the target track is clearly detected with the CDP system. At the single ping level there are pings that exceed the detection threshold with the ADP but are marginal, or undetectable, with the CDP.

4.8 ADP Gains

The gain of ADP relative to CDP can be measured as the difference between the SE's along the target track. These values of gain have been reordered as a function of CDP SE, as shown in Figure 7. It can be seen that the ADP, like most adaptive processors, tends to exhibit largest gains when there is no need for additional gains over a conventional processor. The significant gains of the processor should be measured at threshold signals, those for which the SE for the CDP is near 0 dB. The straight line is a fit to the data, and we can see that at threshold levels the ADP provides a gain in the order of 4-5 dB. There is also a regime from about −5dB to 0dB where the SE is positive only for the ADP method.

5. SYSTEM APPLICATIONS

The ADP method is potentially applicable to both tactical and surveillance sonar systems, is suitable for both monostatic and multistatic configurations, and can be used with most Doppler-sensitive waveforms, such as PCW, FSK, Newhall, or Cox [5]. This algorithm works on one ping at a time, thus providing its gain independent

of that which might result from multiple pings. The algorithm is designed for use when bottom reverberation is a dominant source of noise, so that littoral and/or shallow water operations are most likely to benefit. The algorithm is aimed at extending Doppler sensitivity downward into the one-to-five knot range which is usually obscured, in shallow water, by the zero-Doppler ridge. The algorithm is likely to degrade gracefully with increasing frequency and sea state: it is expected to be useful for most sonar frequencies between ~50Hz and ~1kHz. The ADP algorithm would be expected to work even better in combination with adaptive beamforming, thus forming a spatial/temporal adaptive processing system.

6. SUMMARY

An adaptive Doppler processing algorithm has been described which uses low-rank data matrices both to deal with the non-stationarity of the reverberation background and to permit efficient computation via the SVD. The algorithm has also been desensitized to Doppler replica mismatch, making it robust throughout the entire Doppler spectrum. Performance improvement using at-sea data from the SWAC-3 experiment of ADP versus CDP has been estimated to be near 4-5dB at the threshold of positive signal excess, a significant gain for this data set. It is inferred that this performance gain would be even more important for tracking and classification functions when the actual target Doppler is smaller (absolute value) than the six knots provided by this data set.

7. REFERENCES

[1] James Alsup, "Bistatic Adaptive Processing," *ASW Surveillance Programs—FY97 Program Summaries*, Office of Naval Research, ONR-321-98-3, December 1997, pp.A4/1-7.

[2] Henry Cox and Richard Pitre, "Robust DMR and Multi-Rate Adaptive Beamforming", *Record of the 31st Conference on Signals, Systems & Computers*, Pacific Grove, CA Nov. 1997

[3] D'Amico, A. and X. Zabal, "Adaptive Doppler Detection in Shallow Water," NRaD Technical Report, Code D712, San Diego, CA, 11 December 1995.

[4] James Alsup and Xavier Zabal, "Active Sonar ABF Application Issues", *Record of the 29th Conference on Signals, Systems & Computers*, Pacific Grove, CA Nov. 1995, pp.1081-1085.

[5] Henry Cox, "Geometric Comb Waveforms for Reverberation Suppression," *Conference Record of the Twenty-Eighth Asilomar Conference on Signals, Systems & Computers*, Pacific Grove, CA, Oct.30-Nov.4 1994, pp. 1185-1189.

A Fresh Look At 'Broadband' Passive Sonar Processing

Robert E. Zarnich

Program Executive Office for Undersea Warfare, Advanced Systems and Technology Office,
PEO(USW) ASTO
Passive Processing Project Officer
2531 Jefferson Davis Hwy
Arlington, VA 22242-5169
zarnichre@navsea.navy.mil

ABSTRACT

Traditional passive broadband sonar processing has been constructed by performing spatial decomposition (beamforming) followed by a temporal filter (Eckart weighting) structured to maximize deflection for signal present while minimizing deflection under noise only conditions. This processing, while approaching optimality in the stationary single signal-stationary noise case, is not well suited for clutter rich environments and spectrally diverse signal sets. Techniques such as the Smoothed Coherent Transform (SCOT) and the PHAse Transform (PHAT) as well as a new ad hoc construction referred to as Sub-band Peak Energy Detection or SPED are finding increased usage in modern sonar systems.

This paper discusses sources of the performance gains over traditional processing. A conclusion that 'Passive Broadband' is fundamentally a Direction of Arrival (DOA) estimation problem is developed from this analysis. Based on this conclusion higher resolution techniques including, Quadratic Spectral Capon and Spectral MUSIC are applied to generate a similar broadband output. A comparison of performance to the Navy's current Broadband schemes against a common real data set is presented.

1. INTRODUCTION

The submarine is the primary platform from which the Navy conducts Anti-Submarine Warfare (ASW). In that capacity, and in general, the submarine's greatest ally is stealth. Staying submerged and passively observing the acoustic radiated energy of the surrounding contacts is the preferred mode of operation for the submarine. The Sonar is the eyes of the submarine. In ASW operations in particular, a submarine engaging another submarine, there are no means other than the sonar to observe the target and surrounding area while remaining submerged and preserving stealth.

If the sonar is the eyes of the submarine, 'Broadband' is the peripheral vision. It provides what many refer to as 'Situational Awareness' and others refer to as 'Context'. The Broadband display is one that is used by many operators from the Sonar Operator all the way up to the Commanding Officer. It is a display that provides a relative bearing versus time history of the contacts around the sensor. This is important information since it is the only display that immediately provides a reference of the presence of objects and information regarding their relative motion.

Since the introduction of digital sonars, there have been two approaches to broadband in use for over 20 years. They are a traditional square law, or energy detector process, and a split aperture cross correlator. In the past several years two techniques have extended these basic concepts and recently found use in the Navy's latest Submarine Sonar development. Correlation processing was enhanced via a whitening process developed by Carter et.al. [1], the Smoothed Coherence Transform (SCOT). The energy detector was enhanced with a non-linear peak picking operation for each frequency band over azimuth prior to integration over frequency. This technique is named the Sub-band Peak Energy Detector (SPED) and was original developed by Hughes Aircraft Corporation in Fullerton CA. Although the later of these techniques is an ad hoc approach, upon further analysis one discovers that its roots are in a well researched field, Direction Finding (DF). Based on this discovery application of a basic direction finding technique, Quadratic Capon Spatial Spectral Estimation, (MPDR), to the power accumulation concept of SPED yield interesting results.

2. Passive Acoustic 'Broadband'

In the functional capacity described here, it would be expected that all contacts would be presented on the display regardless of the type of contact. The azimuthal information is important for surfacing a submarine to periscope depth and therefore screening out fishing boats

in favor of submarines may not have a desirable effect. Desired attributes of a Broadband Processor might be best defined as:

1) A summary of detectable contacts

2) Precise location in time and Azimuth

3) Intuitive Format

4) Detail not Clutter

For the purposes of this paper we will define Clutter as a contact the masks another otherwise detectable contact on the broadband display. Detail would be best described as high resolution.

2.1 Traditional Energy Detection Based Broadband

Figure (1) outlines the general approach to the second-generation passive broadband. It is refered to as second generation since the first generation broadband was processed on an analog system and this approach was subsequently realized on a digital system, hence second generation. Beamforming is either conventional or adaptive. (Although in the first and fielded submarine sonars it was a conventional time delay and sum approach.) The beams are steered such that the beams overlap at approximately their half power. Spectral Weighting is comprised of time to frequency transformation, magnitude squaring, and a weight chosen as prescribed in Eckart [2]. Spatial Normalization is a split window normalizer. This process scales measured spatial power value to a normalized distribution expected by the requantization process. The Non-Linear Requantization maps the power values to a reduced dynamic range while retaining a marking density on the display that is sensitive to levels above expected noise while not being over driven by strong signals. The mapping is chosen based on empirical historic data.

This processing as discussed in [2] and other classic works is developed for single signal detection, of known spectral content, in the presence of a known noise field. Unfortunately the noise field is neither fixed in time or in space over any large time interval. Based on experience and empirical data the presumption that the signal and noise spectra is known and fixed is not well founded. To realize this consider the earlier observations that observing surface craft is equally as important under different operating conditions as detecting and tracking submarines during ASW operations. Therefore, although the process is developed as an optimal one, it is ill posed for the requirements of a Broadband Processor as outlined in 2.0.
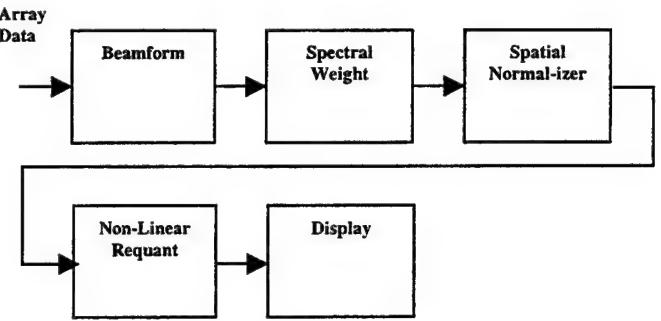

Figure(1) Traditional Energy Detection (Square Law) Broadband Block Diagram

2.2 Current Broadband

In the past decade with the advances in computing resources available to implement sonars more elaborate schemes have been formulated and two have found success. The SCOT whitened split aperture correlator and the SPED algorithms as discussed earlier are two of those. Both of these algorithms have matured through application on various arrays over a variety of frequency bands. The correlation based broadband provides best overall performance during periods of physical non-stationarity such as that seen during own ship turns. SPED provides support for detecting contacts whose broadband signature may not be at minimum detectable levels but contains significant power in temporally narrow spectral components. In this way the two techniques are complementary since the whitening component of the correlation process eliminates these narrow contributions, and the necessary spatial-temporal normalization for proper operation of SPED is difficult during the periods which lack physical stationarity.

2.2.1 Correlation Based Broadband

Figure(2) provides an outline to a generalized correlation based broadband. Of interest is the striking common ground to direction finding. The aperture is separated into two equal sub-apertures. In the current system a beam level adaptive beamformer with 7 degrees of freedom, with a unity gain constraint, a soft mainlobe pattern constraint, diagonal loading, and a weight norm adjustment is employed. Beams are formed with overlap on the order of the 1 dB down intersection. The SCOT whitening algorithm is used [1], and beams with common steering directions are cross correlated with a lead, lag, and no common azimuth delay terms. The resulting correlation scan is up-sampled to smooth the displayed scan and a similar quantization process used.
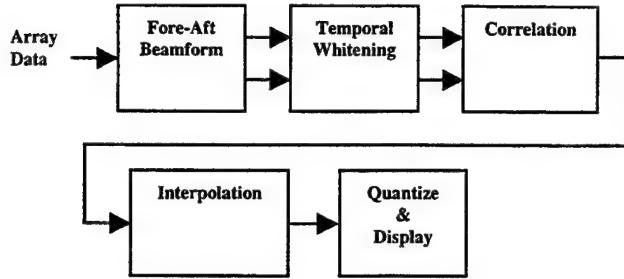
Figure(2) General Correlation Broadband Block Diagram

MacDonald and Schultheiss [3] showed that a split aperture cross-correlation processor was an optimal method for direction finding with broadband contacts under a limited set of conditions. This is an important observation when researching enhanced broadband processing.

2.2.2 Peak Energy Detection

Figure(3) provides an outline to the generalized SPED processing concept. The beamforming has the same characteristics as the beamformer discussed in 2.2.1 however the entire aperture is used. The second stage, Normalization and Peak Picking, de-trends the data across frequency. Sea noise has power greatest at lower frequencies, decreases on the order of 8-10dB per octave over 1-10Hz, flattens out over the 10-100 region and decreases 5-6 dB per octave over the 100-100k Hz region for deep-water spectra. Deep water and other noise measurements are discussed in detail in Urick [4] and Burdic [5] along with many others. The de-trending is to prevent any particularly loud band from dominating the accumulated power. The peak picking operation is done over azimuth for each frequency. The maximum number of observable peaks in any given scan is given by the simple expression in equation(1). This is due to the polynomial behavior of the array manifold vector $\mathbf{v}_s$ when evaluating the spatial spectrum. Recall:

$$\mathbf{v}_s = e^{-j 2\pi \frac{\mathbf{d}}{\lambda_s} \cos(\theta_s)}$$

and $\mathbf{d}$ is an Nx1 vector with the values of the position of the element along the axis of the array, λ_s is the wavelength of the signal of interest, and θ_s is the angle of the signal path intersection with the axis of the array as measured from the forward end of the array. When considering the outer product of the sampled data and the steering vector the spatial spectrum has a limited number of modes. The number of wavelengths over the aperture dictates the possible number of observable peaks as defined in equation (1).

$$Peaks = \left(N * \frac{f_{signal}}{f_{design}} \right) - 1 \qquad (1)$$

Peak picking identifies the beam and therefore steering angle closest to the true direction of arrival. Fine Bearing Interpolation is done to register the peak in a azimuth grid finer than the one used to create the original beam set. Assuming approximately linear beam power versus azimuth, since beams are highly overlapped, a direction of arrival (DOA) is estimated from the powers of the highest beam and the two adjacent beams and the cell in the grid closest to the DOA estimate is retained as best estimate. Accumulation over frequency is a summing operation. In the current implementation there are two modes. The first retains only the presence of the peak, meaning a one is placed in the cell. This mode is referred to as 'Clutter Suppress'. The second places the power estimate in the cell in a fashion consistent with traditional energy detection. Both modes are summed over frequency to create a scan in time. Spatial Normalization is similar to that discussed in 2.1 as is the quantization process. The display is than rendered and updated for various integration intervals.
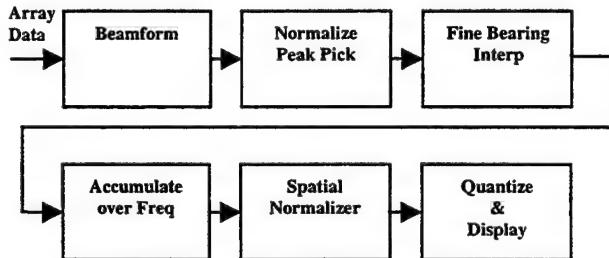
Figure(3) General SPED Broadband Block Diagram

2.3 Observations on Current Techniques.

These techniques were developed through extensive use of real data from fixed and moving arrays. They have proven quite successful. None of the techniques incorporate a frequency-weighting scheme other than the cut off edges. In current implementations there are multiple bands running concurrently. In both methods a common theme of increased precision of DOA estimation exists. The whitened correlation process is optimal under limited conditions for broadband as discussed in [3]. A beam scan peak pick is also optimal for narrowband under limited conditions as discussed in [5]. As alluded to in section 2.0, resolution and acuity are the key elements in making the broadband processing successful when the temporal filtering is an all pass filter.

Since the essence of these processes appears to be the ability to perform direction finding combined with novel approaches to accumulate and summarize the contacts in an intuitive display format it begs the question, what would better Direction Finding provide?

3.0 Direction Finding Based Broadband

Figure(4) represents a general approach to exploit a DF processor followed by a power accumulation scheme modeled after SPED. Van Trees [6] details various approaches to doing the Covariance Estimation as well as the Direction Finding. The classic MVDR result when working with an unknown covariance becomes Minimum Power Distortionless Response (MPDR) and can be used as a direction finding technique. When one evaluates the reciprocal of the quadratic product of steering vector and the inverted Covariance Estimate versus scan bearing, θ_{scan}, equation (2), the value is proportional to power with a scaling factor of $1/N^2$. This then corresponds to a spatial spectral estimate for the sample period over which the covariance, R_x, is estimated. When only the Noise Sub-Space eigenvectors are used the familiar MUSIC algorithm results, however the value is no longer proportional to power.

$$Power(\theta) \propto \frac{1}{\mathbf{v}^H(\theta)\mathbf{R}_x^{-1}\mathbf{v}(\theta)} \qquad (2)$$

In the form presented [6] refers to this as a member of the Quadratic Spectral Algorithms, specifically the Quadratic Spectral Capon Algorithm. The algorithm estimates the DOA's as those which maximize Power(θ_{scan}). One could choose to invoke a threshold if screening against a minimum SNR. Similarly MUSIC retains the values which maximize the expression for the noise sub-space. The Accumulation over frequency, Spatial Normalization, and Quantization and Display will follow similarly from the SPED outline in 2.2.1
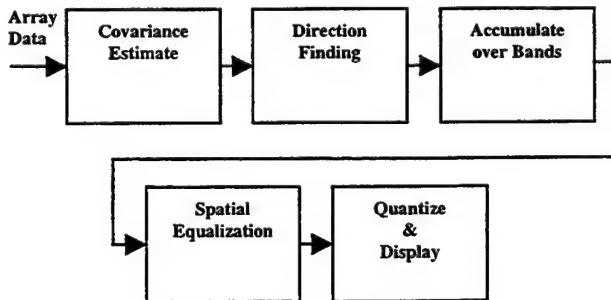
Figure(4) General DF Based Broadband Block Diagram

3.1 Prototype DF Based Broadband Implementation

The goal was to construct a DF based approach to Broadband to enhance overall performance while retaining the ability to process through own-ship maneuvers with the capability of the correlator and retain the sensitivity to strong narrowband contacts as in SPED. The design must support the update intervals required of a tactical system. This really means that we can't simply integrate over any period of our choice since information must be updated frequently to be relevant. Initially, integration periods for R_x were chosen consistent with the integration periods used for the current SPED process. The processing was done on a 48 element linear array. Specific bands of interest and resolutions would make this discussion classified and are omitted.

3.1.1 Estimate the Covariance

The integration time, 5 seconds, chosen initially did not provided enough sample support to guarantee a positive definite estimate so diagonal loading was employed. As a good rule of thumb, [6] suggests that loading greater than the signal of interest and on the order of 10dB less than the loudest interferer provides robust performance. This was used as a starting point and through experimentation settled on in the prototype implementation. To do this R_x is normalized by its trace divided by its rank. This leads to an expected value of the element variance of 1, to which -10dB times an identity matrix is added to the normalized R_x. The normalization term is retained as it is needed to estimate the true power along DOA estimates. Eventually an exponentiation time of 20 seconds was arrived at to provide robust performance.

3.1.2 Direction Finding

The Quadratic Spectral Capon (QSC) algorithm and MUSIC were both used. The number of steering directions, over which spatial power estimates were made, chosen was 10*N+1. This was reduced by a factor of 2 however the visual effects on the display were not appealing so 481 directions were retained. Many attempts to choose a fixed number of degrees of freedom for MUSIC were made but the results were mixed. At times MUSIC appeared better than QSC but the goal was robust performance so QSC is recommended for this application.

Many effects plague the sub-space techniques such as correlated paths (Multi-Path) which are quite normal in undersea acoustics. Using AIC and MDL for model order estimation dynamically proved to provide poor performance. Since SNRs of targets of interest tend to be low these tests are not very effective. Until better techniques for model order estimation are available it

would appear sub-space approaches for DF would not be viable for problems where the model order is highly variable and signals of interest are small, such as this problem.

3.1.3 Accumulate Over Frequency

Both methods discussed in 2.2.2 were used. Almost uniformly the power summing method dominated the performance of the binary approach. This is the recommended method.

3.1.4 Spatial Normalization

Although the results indicate that some method of spatial normalization would improve performance, none was done in the interest of time. The normalization process only enhanced an already discernable contact but added a factor of about 15 to the execution time of the prototype algorithm.

3.1.5 Quantization and Display

The accumulated power was converted to 'dB', using 10*log of the spatial power estimate. MATLAB™ *imagesc* and *image* commands displays were used to render the results to image shown in 5.d.

3.4 Results

Figures (5.a), (5.b) and (5.c) are the SPED, SPED CS and Correlation Broadbands. Figure (5.d) is the 20 second averaged R_x QSC algorithm with power accumulation. Of note is the maintenance of performance through turns, suppression of broadside array noise, and the overall sharpness of the traces. This is just one data set and many more have been used to test the concept. One can see quite clearly the improvement in overall performance in these graphics. The processing is also providing very precise location for some isolated narrowband components that are otherwise collected in weak traces on the SPED broadband.

4.0 CONCLUSIONS

Performance enhancements are available through improved Direction Finding based processing. This study indicates there still may be significant room for improvement however more research is required. Several areas to be pursued include the following. A capacity for better subspace estimation would allow a more robust employment of MUSIC as a DF technique for improved closely spaced contact resolution. Integration time extension during periods where the objects retain temporal stationarity, but lack physical (orientation) stationarity is

an area which appears to have interesting implications. If longer integration times can be had without the smearing effects in the Covariance, contacts can be detected at reduced SNRs. Additionally resolution can be enhanced through reduced bearing variances, thereby reducing overlap of closely spaced traces on the display. Colorization schemes and other approaches could be exploited to improve the visual presentation as some were explored but are beyond the bounds of this paper.

4   REFERENCES

[1] Carter G. C., Nuttal, A. H., and Cable P. G., "The Smoothed Coherence Transform", *Proc.IEEE(Lettl)*, 61, 1497, 1973

[2] Eckart C., "Optimal Recifier Systems for the dectection of steady signals', University of California, Scripps Inst. Oceanography, San Diego, CA, Marine Physical Lab. Rep s10 12692, S10 Ref 52-11, 1952

[3] MacDonald, V. H. and and Schultiess, P. M., "Optimum Passive Bearing Estimation", *J. Acoust. Soc. Am.*, 46, 37, 1969.

[4] Urick, R. J., *"Principles of Underwater Sound"*, 3ed, Reprint, 1983, Peninsula Publishing

[5] Burdic, W. S. *"Underwater Acoustic System Analysis"*, 1984, Prentice Hall

[6] Van Trees, H. L., *"Detection Estimation Theory Volume 4: Optimal Array Processing"*, 1998 Manuscript, George Mason University.

Figure 5.a SPED Clutter Suppress (SPED-CS)
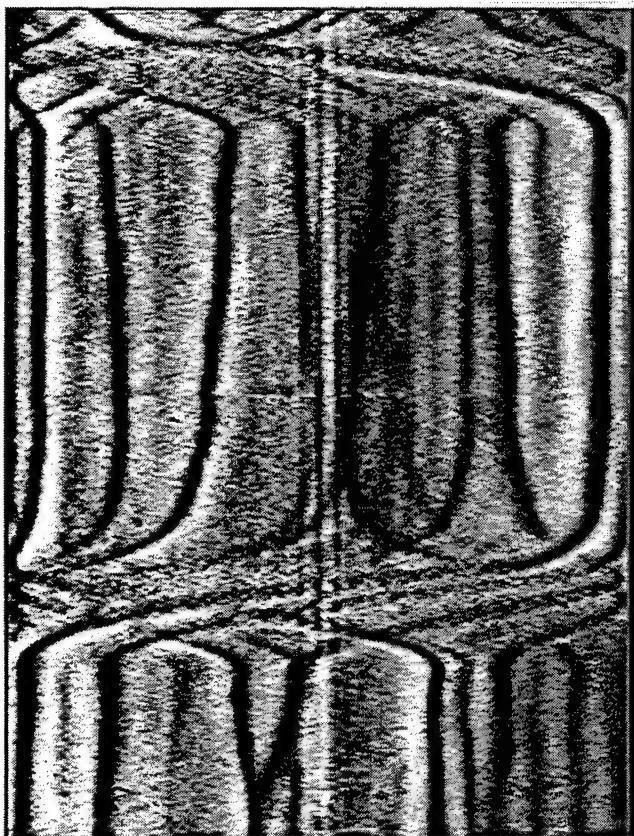

Figure 5.b SCOT Correlation
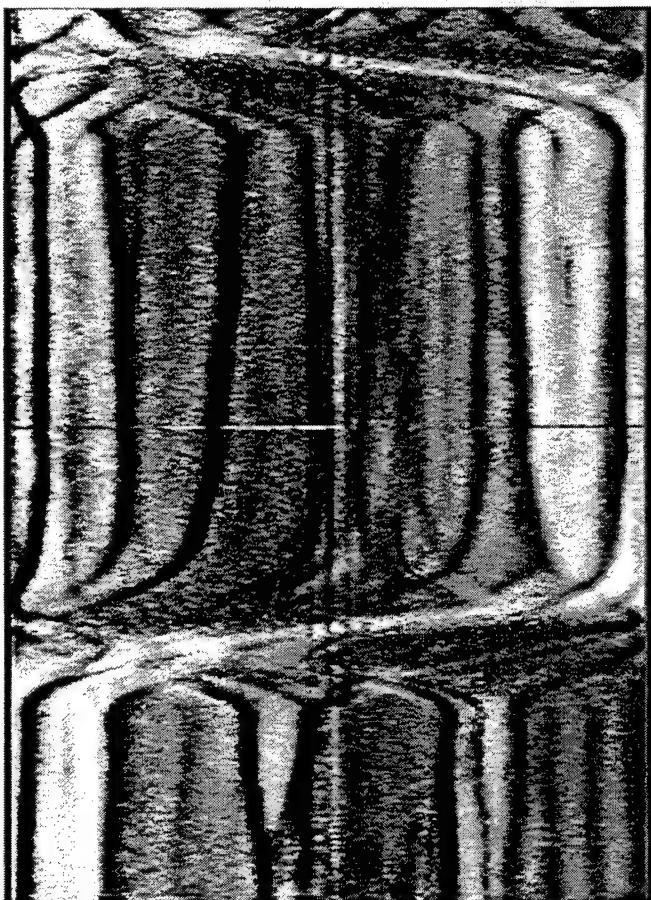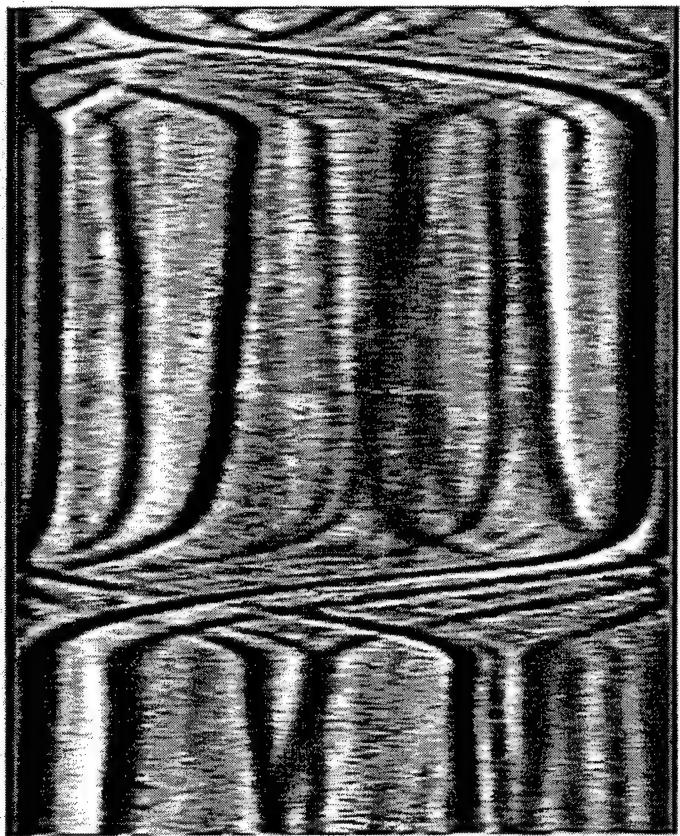
Figure 5.c SPED Energy Detection
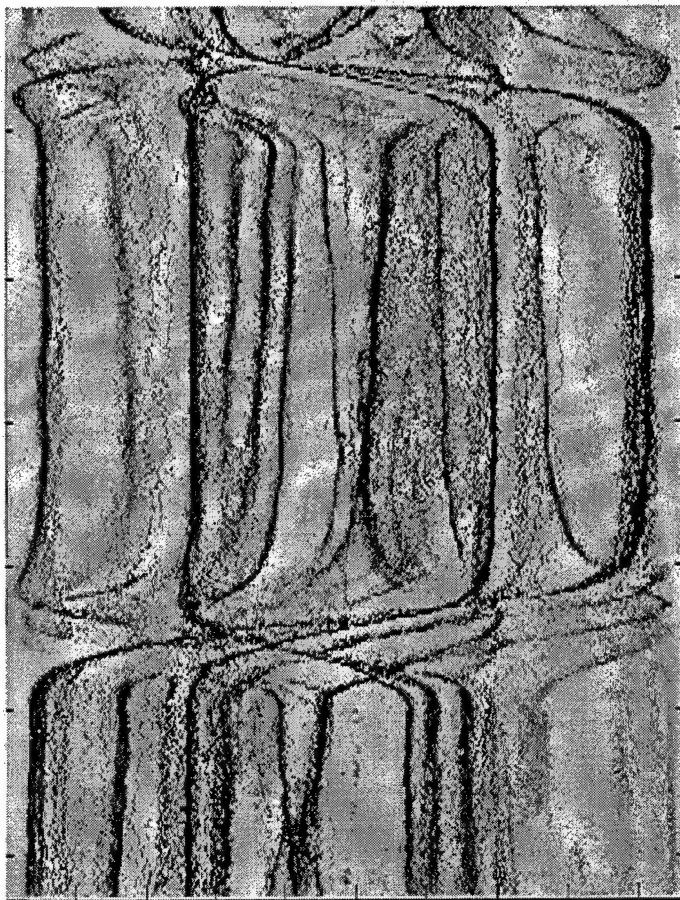
Figure 5.d 20 Second Rx DF Based Broadband

104

Over-the-Horizon Skywave Radar Target Localization

Richard Anderson, Jeffrey Krolik, and Michael Papazoglou

Duke University

Department of Electrical and Computer Engineering

Durham, NC 27708

Abstract— **Target localization using over-the-horizon skywave radar relies critically on correctly modeling and processing multipath returns. This paper deals with the use of both unresolved and resolved multipath arrivals in order to perform 3-D estimation of aircraft position. In particular, unresolved micro-multipath returns due to surface reflections local to the target are used for altitude estimation and the combination of resolvable multipath arrivals from different layers of the ionosphere are used to enhance target latitude and longitude estimates. For altitude estimation, the maximum likelihood method presented here exploits a statistical model for the altitude-dependent complex fading of the target across radar revisits. For latitude and longitude determination, the proposed maximum a posteriori mode linker uses a statistical model of the ionosphere to achieve robust localization in the presence of uncertain channel characteristics. Real data results are presented for both methods which demonstrate that altitude can be estimated to within 3000 ft. and as much as a 3:1 improvement in ground coordinate accuracy can achieved.**

I. INTRODUCTION

Multipath propagation is a dominant feature of skywave high frequency (HF) electromagnetic propagation through the ionosphere. In particular, refraction of over-the-horizon (OTH) HF signals through different layers of the ionosphere often result in resolvable target returns in delay and Doppler. In principle, when the propagation paths taken by these arrivals are correctly identified and further these returns are properly associated with the same target, the complexity of the multipath channel can be exploited to improve target ground coordinate estimation. However, this so-called mode linking process can be very sensitive to detailed knowledge of the downrange ionosphere. In section III, a statistical model of ionospheric uncertainties is employed to obtain a maximum a posteriori mode linker which is inherently more robust to errors in the ionospheric model.

Another example of multipath in OTH are unresolvable micro-multipath returns from the direct and surface reflected arrivals of an aircraft target. Typically these "micro-multipath" returns are not resolvable in delay or Doppler, however, their interference causes complex fading from dwell-to-dwell which is characteristic of aircraft altitude. In section II, the concept of matched-field processing (MFP), borrowed from passive sonar, is used to develop a maximum likelihood estimate of altitude using multiple radar dwells.

II. MATCHED FIELD ALTITUDE ESTIMATION

OTH radars are currently capable of localizing targets in ground range and azimuth, [1], [2], [3], but altitude estimation has thusfar not been reliably achieved. In this section, a matched-field estimate of altitude which uses multiple radar dwells is presented. Although a large literature on matched-field processing (MFP) techniques exists for underwater source localization [4], the approach is relatively new in radar applications. In a general sense, MFP consists of correlating the received data with predictions from multipath propagation models for a set of hypothesized target locations. In previous work, electromagnetic MFP has been applied to low angle line-of-sight radar for height finding in the presence of specular multipath reflections from the ground surface [5] and the estimation of tropospheric refractivity parameters using point-to-point microwave transmissions [6]. It has also recently been proposed for single-dwell target altitude estimation with OTH radar [7]. MFP for altitude estimation exploits the altitude dependence of differential delays and Dopplers between the micro-multipath returns from a bistatic skywave radar illustrated in Figure 1. Although this approach exploits the altitude dependence of unresolved multipath returns in complex delay-Doppler space, its performance suffers in situations where the coherent integration time (CIT) of the radar is short.

The multiple-dwell matched-field altitude estimation technique proposed here exploits the altitude dependence of dwell-to-dwell shape changes in the complex delay-Doppler multipath return. To handle slowly changing amplitude variations due to Faraday rotation and aspect-dependent target backscatter characteristics which are not target altitude dependent, a Markov model for the multipath reflection coefficients is used. This leads to a time-evolving maximum likelihood (ML) estimate of altitude which is derived in Section 2b. Monte Carlo simulations and a result with real data presented in Section 2c suggests that using using short CIT's, moderate signal bandwidth, and a 30 second revisit rate, multi-dwell matched-field estimation can achieve better than 5,000 ft. accuracy after as few as four radar dwells.

A. Modeling Micro-Multipath Radar Returns

The signal model described here is based on an FM/CW radar system which is able to extract both time delay and Doppler information from a target. For the kth revisit, the radar transmits a coherent series of linear FM chirp waveforms with waveform repetition interval $T(k)$ and sweep rate $b(k)$. The reflected signal contains contributions from the $L = 4$ micro-multipath ray combinations, shown in figure 1. Slant range is estimated by performing a DFT over a pulse repetition period. A second DFT is performed over several pulses to obtain Doppler shift.
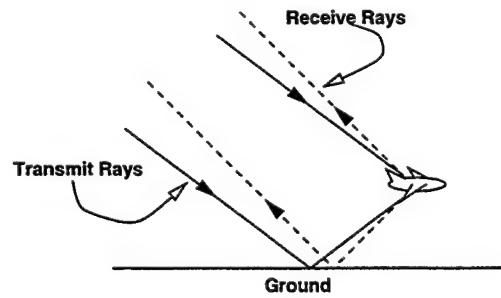

Fig. 1. Multipath raypaths local to the aircraft target for transmission and reflection with a bistatic radar.

Consider the radar return due to a single bistatic ray path, l. This component of the return will have a time delay $\tau_l(k)$ and a Doppler shift $\omega_l(k)$ which are a function of aircraft altitude, z, as well as the slant range, $g_0(k)$, and Doppler, $\omega_0(k)$, of a

ray which intersects the ground directly beneath the aircraft. For slant range index n, Doppler index m, and time $t(k)$, let the post-DFT slant range waveform for the lth component be $R_l[n, k]$ and the post-DFT Doppler waveform be $P_l[m, k]$. The model for the complex range-Doppler return is

$$x[n, m, k] = \qquad\qquad\qquad\qquad\qquad\qquad (1)$$
$$e^{j\theta_k} \sum_{l=1}^{L} c_l(k) R_l[n, k] P_l[m, k] e^{j\omega_l t(k)} + \eta[n, m, k]$$

where $c_l(k)$ denotes the complex amplitude of the lth micro-multipath ray, θ_k is the unknown starting phase of the dwell, and $\eta[n, m, k]$ represents additive noise.

Let the $NM \times 1$ vector $\mathbf{x}_k$ represent an $N \times M$ block of the complex range-Doppler map in the neighborhood around the slant range and Doppler of the target of interest. The data model in (1) can then be written as

$$\mathbf{x}_k = e^{j\theta_k} \mathbf{H}_k \mathbf{D}_k \mathbf{c}_k + \mathbf{n}_k \qquad\qquad (2)$$

where

$$(\mathbf{H}_k)_{nN+m,l} = R_l[n, k] P_l[m, k] \qquad\qquad (3)$$

and $\mathbf{D}_k$ is a diagonal matrix with $(\mathbf{D}_k)_{l,l} = e^{j\omega_l(k)t(k)}$. The complex micro-multipath ray amplitudes are treated here as zero mean Gaussian random variables with covariance

$$E\left[\mathbf{c}_k \mathbf{c}_k^H\right] = \Lambda_c = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_L \end{bmatrix}, \qquad (4)$$

and the noise is assumed to be zero mean Gaussian with covariance $\sigma_n^2 \mathbf{I}$. Thus $p(\mathbf{x}_k|z)$ is Gaussian with mean zero and covariance $\mathbf{R}_{k,k} = \mathbf{H}_k \Lambda_c \mathbf{H}_k^H + \sigma_n^2 \mathbf{I}$.

B. Maximum Likelihood Estimation with Multiple Dwells

Let the set of data snapshots for the $K+1$ revisits, $k = 0 \ldots K$ be denoted by $\mathbf{X}_K$, and the associated phase differences by $\Delta\Theta_K = (\Delta\theta_1, \Delta\theta_2, \ldots, \Delta\theta_K)$, where $\Delta\theta_k = \theta_k - \theta_{k-1}$. In order to obtain a multiple dwell ML estimate of altitude, the joint distribution of $\mathbf{X}_K$ conditioned on $\Delta\Theta_K$ and z is needed. In the absence of Faraday rotation and aspect dependent target backscatter characteristics, the complex micro-multipath ray amplitudes should be perfectly correlated so that $E\left[\mathbf{c}_k \mathbf{c}_{k-\mu}^H\right] = \Lambda_c$ for arbitrary μ. In practice, slow random fluctuations can be handled by modeling $\mathbf{c}_k$ as a first-order Markov process, in which case

$$p(\mathbf{X}_K|z, \Delta\Theta_K) = p(\mathbf{x}_0|z) \prod_{k=1}^{K} p(\mathbf{x}_k|\mathbf{x}_{k-1}, \Delta\theta_k, z). \qquad (5)$$

In general the term $p(\mathbf{x}_k|\mathbf{x}_{k-1}, z)$ describes the variation of the signal from one dwell to the next.

An ML estimate of altitude is obtained by maximizing $p(\mathbf{X}_K|z, \Delta\Theta_K)$ with respect to $\Delta\Theta_K$ and z. Assuming that consecutive micro-multipath amplitudes $\mathbf{c}_{k-1}$ and $\mathbf{c}_k$ are perfectly coherent, then $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \Delta\theta_k, z)$ is also Gaussian with mean and covariance

$$\mathbf{m}_k = \mathbf{R}_{k,k-1} \mathbf{R}_{k-1,k-1}^{-1} \mathbf{x}_{k-1} \qquad\qquad (6)$$
$$\mathbf{Q}_k = \mathbf{R}_{k,k} - \mathbf{R}_{k,k-1} \mathbf{R}_{k-1,k-1}^{-1} \mathbf{R}_{k,k-1}^H$$

where $\mathbf{R}_{k,k-1} = \mathbf{H}_k \mathbf{D}_k \Lambda_c \mathbf{D}_{k-1}^H \mathbf{H}_{k-1}^H e^{j\Delta\theta_k}$. Maximizing the log of the likelihood in (5) over $\Delta\Theta_k$ yields

$$L(z) = \log p(\mathbf{x}_0|z) + \sum_{k=1}^{K} L_k(z) \qquad\qquad (7)$$

where $L_k(z) = \max_{\Delta\theta_k} \log p(\mathbf{x}_k|\mathbf{x}_{k-1}, z, \Delta\theta_k)$, which, from (6), can be written

$$L_k(z) = -\log \pi^N |\mathbf{Q}_k| - \mathbf{x}_k^H \mathbf{Q}_k^{-1} \mathbf{x}_k \qquad (8)$$
$$+2|\mathbf{x}_k^H \mathbf{Q}_k^{-1} \mathbf{P}_k \mathbf{x}_{k-1}| - \mathbf{x}_{k-1}^H \mathbf{P}_k^H \mathbf{Q}_k^{-1} \mathbf{P}_k \mathbf{x}_{k-1}$$

where $\mathbf{P}_k = \mathbf{H}_k \mathbf{D}_k \Lambda_c \mathbf{D}_{k-1}^H \mathbf{H}_{k-1}^H \mathbf{R}_{k-1,k-1}^{-1}$. The ML estimate of altitude can now be obtained by a one dimensional numerical maximization of (7) with respect to z, which is computationally efficient.

Note that the coherence assumption of $\mathbf{c}_k$ is limited by effects such as Faraday rotation. Over long target tracks, Faraday rotation will have the effect of decorrelating the complex ray amplitudes. If the Faraday decorrelation is known, it can be incorporated into $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \Delta\theta_k, z)$. It should also be noted that the likelihood functions in this section were derived assuming that $g_0(k)$ and $\omega_0(k)$ are known. In practice, these can be estimated by finding the peak correlation of $\mathbf{x}_k$ with the signal model in (3) for the strongest radar returns and interpolating at the weaker revisit times along the radar track.

C. Simulation and Real Data Results

To evaluate the expected performance of the ML altitude estimation approach presented here, the probability of correct localization (PCL) within a 5,000 ft. altitude band is estimated over 200 Monte Carlo simulations as a function of SNR and target altitude. The CIT is nominally 2.5 seconds, the radar bandwidth is 17 kHz, and the operating frequency is 10 MHz. The target ground range is 1200 km and the radial velocity is -190 m/s. The revisit interval between dwells is 30 seconds. The ionosphere was modeled using a single quasi-parabolic E-layer [8] with critical frequency 3.5 MHz, height 110 km, and thickness 32 km.
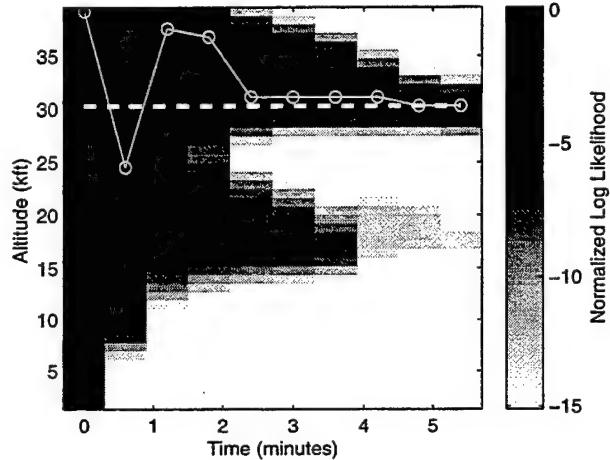

Fig. 2. Simulation log-likelihood function evolving over 30 second revisits. True altitude is 30,000 ft., and estimated altitude is 30,000 ft. after 5 minutes (9 dwells). The SNR is 25 dB, the nominal CIT is 2.5 seconds, and the bandwidth is 15 kHz.

Figure 2 illustrates an example of the log likelihood function in (7) evolving over five minutes for a target altitude of 30,000 ft. The bandwidth and CIT are clearly not sufficient to estimate altitude with a single dwell, but after five dwells the ML estimate converges to the correct altitude. Figure 3 shows the PCL for a fixed altitude of 26,000 ft. as a function of time for varying SNR over 200 random trials. After ten minutes, the PCL reaches 0.8 for a 15 dB target. However, for SNR's of 25 dB and above, this level of performance is achieved after
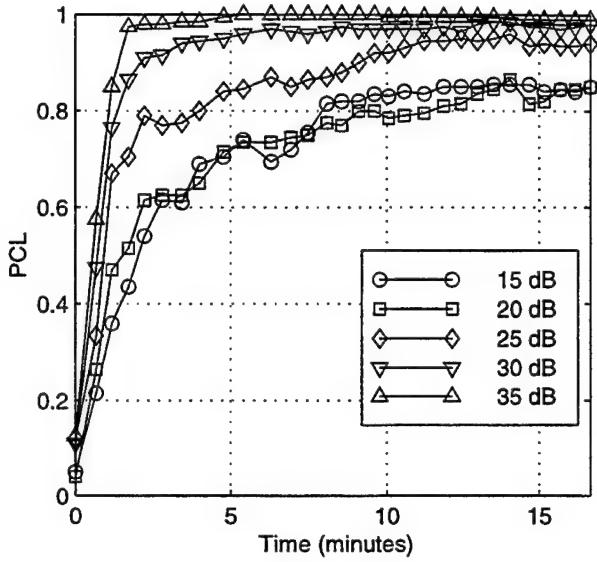
Fig. 3. Probability of correct localization (PCL) within 5,000 ft. bands vs. time for a 26,000 ft. target for varying SNR (200 realizations).

only four minutes. These results indicate that the performance threshold for this estimator is near 20 dB.
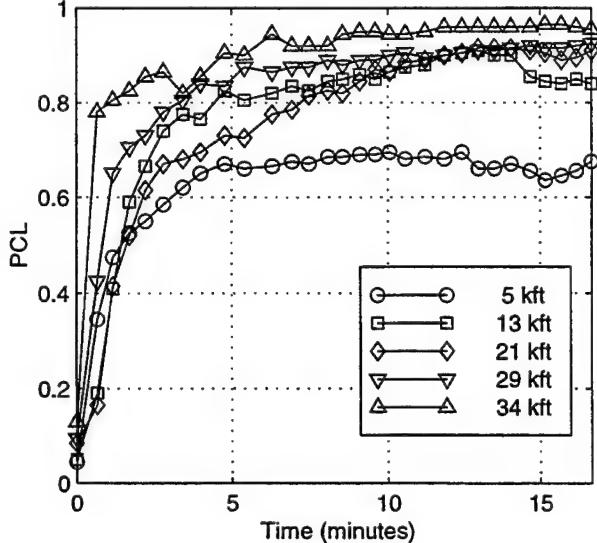

Fig. 4. Probability of correct localization (PCL) within 5,000 ft. bands vs. time for a 25 dB target at different altitudes (200 realizations).

Figure 4 shows the PCL results as a function of target altitude for a fixed SNR of 25 dB. In general, estimation performance improves with increased target altitude. Note that the range-Doppler map of the received radar signal is composed of L returns due to the ray combinations described in section II-A. As the target altitude increases, the separation of these returns in the range-Doppler map also increases. This increased separation might be expected to lead to increased accuracy in the altitude estimate.

Figure 5 shows the result from processing real data from an OTH radar track of a strong, high altitude commercial aircraft. The resulting altitude estimate in Figure 5 converges to within
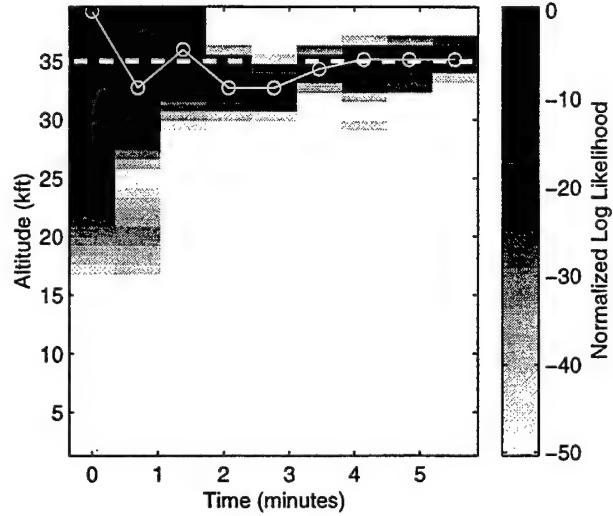

Fig. 5. Real data log-likelihood vs. time, true altitude is 35,000 ft., estimated altitude is 35,200 ft.

200 ft. of the true aircraft altitude after four minutes. Furthermore, the evolving likelihood function shows very similar behavior to the simulation likelihood function in Figure 2, which serves to validate the simulation performance described above.

III. Maximum *A Posteriori* Mode Linking

HF radiowave propagation through the ionosphere often gives rise to multiple raymodes between the radar and a target which results in multiple slant tracks from a single target [1], [2], [9]. To resolve the multi-path and multi-target ambiguities for a given ionospheric plasma frequency profile and dwell illumination region (DIR), ray tracing is typically used to determine the coordinate registration (CR) transformations from tracks in slant coordinates to target locations in ground or geographical coordinates. To achieve greater robustness to uncertainty in down-range environmental conditions, the plasma frequency profile parameters are modeled here as random variables with statistics determined from a QVI and WSBI measurements. A statistical model of the OTH radar slant track data, described in Section 3a, can then be formed by mapping the random profile parameters through Monte Carlo evaluation of a ray tracing propagation model. The objective of maximum *a posteriori* (MAP) mode linking, derived in Section 3b, is to find the most likely slant-track-to-target association and most likely target ground locations given the slant track observations and a model for their underlying probability distribution. The implementation of the MAP mode linker and real data results are presented in Section 3c.

A. Modeling Multipath Slant Tracks

In the modeling of the slant track data, raymode elevation angle information can be exploited by ordering the slant tracks according to increasing Doppler frequency magnitude. The Doppler frequency magnitude of a return is proportional to the average of the target's velocity projections onto the transmit and receive raypaths. Since the ray elevation angles undergo only extremely small changes from revisit to revisit, the Doppler ranking provide a robust means of incorporating ray elevation angle information which is consistent from revisit to revisit. At the kth revisit, let $x_{j,k}$ denote the slant range of the jth Doppler-ordered slant track, propagating on raymode $s_{j,k}$ and associated with a target at ground range r_k. The slant range observations

are then modeled by

$$x_{j,k} = F_{s_{j,k}}(r_k) + n_{j,k} \qquad (9)$$

where $F_{s_{j,k}}$ is the ground range to slant range transformation for raymode type $s_{j,k}$ and $n_{j,k}$ is zero-mean Gaussian distributed slant track jitter. Each slant track observation is modeled as a doubly stochastic random variable since a slant range observation's distribution is determined by a) the probability that it corresponds to a particular elevation angle dependent raymode type where the raymodes themselves are random variables and b) the slant range distribution given both the raymode type and the slant tracks with neighboring elevation angles.

B. MLCR and MAP Mode Linking

For an individual target where the slant-track-to-target association is known *a priori*, the raymode assignments and target ground location can be jointly determined by a previously developed method called maximum likelihood coordinate registration (MLCR) [3], [10]. Modeling the slant ranges $\mathbf{x}_k = [x_{1,k}, \ldots, x_{J,k}]^T$ as a Markov process, i.e.

$$p(\mathbf{x}_k|\mathbf{s}_k, r_k) = \prod_{j=1}^{J} p(x_{j,k}|x_{j-1,k}, s_{j,k}, s_{j-1,k}, r_k),$$

and modeling the raymodes, $\mathbf{s}_k = [s_{1,k}, \ldots, s_{J,k}]^T$, as a Markov chain, i.e. $P(\mathbf{s}_k|r_k) = \prod_{j=1}^{J} P(s_{j,k}|s_{j-1,k}, r_k)$, the maximum likelihood (ML) estimate of target ground position is given by

$$[\widehat{r}_k, \widehat{\mathbf{s}}_k] = \arg\max_{r_k, \mathbf{s}_k} \sum_{j=1}^{J} \left(\log p(x_{j,k}|x_{j-1,k}, s_{j,k}, s_{j-1,k}, r_k) \right.$$
$$\left. + \log \Pr(s_{j,k}|s_{j-1,k}) \right) \qquad (10)$$

where the raymode transition probabilities, $\Pr(s_{j,k}|s_{j-1,k})$, and the output probability distribution parameters are estimated via Monte Carlo raytracing. The Markov approximations allow the use of a fast dynamic programming method for obtaining the MLCR ground location and raymode sequence estimates. The transition probabilities model the joint probability of different raymodes to the same location on the ground. The output probabilities model the correlation among the slant tracks due to common reflections heights on the transmit or receive paths. Both terms limit the possible raymode sequences and ground locations to those that are consistent with the magnitude Doppler ordering and the observed slant track data.

In order to perform mode linking in multiple target scenarios, the MLCR method must be evaluated for all possible slant-track-to-target associations. For a set of N_k slant tracks, this may be accomplished by performing a generalized likelihood ratio test (GLRT) to select the best mode linking hypothesis. The possible slant-track-to-target associations may be enumerated by using a mode linking hypothesis tree such as the one shown in figure 6 for 4 slant tracks and 15 hypotheses. From the top "root" node down, each branch corresponds to a different slant-track-to-target assignment to be evaluted. For instance, under H_1, each slant track corresponds to a different target; under H_2, slant tracks 3 and 4 correspond to target 3; and under H_{15} all 4 slant tracks correspond to the same target. Note that the mode linking hypotheses implicitly provide *a priori* relative target ground coordinate information. For example, under H_2 in figure 6, the ground ranges, ground azimuths and ground range-rates of slant tracks 3 and 4 should be very similar if these two slant tracks do originate from the same target. Combining
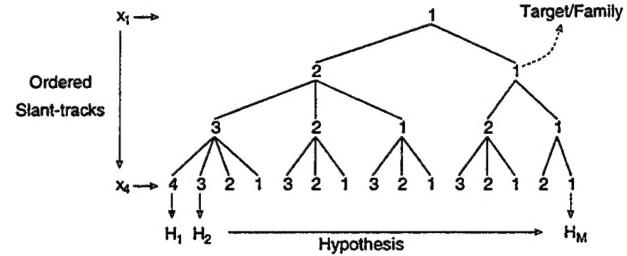

Fig. 6. Mode Linking Hypothesis Tree

the mode linking GLRT with *a priori* ground coordinate distribution, the optimal mode linking hypothesis can then be found by substituting the slant track data, $\mathbf{X}_k$, the MLCR estimates, $\widehat{\mathbf{s}}_{m,k}$ and $\widehat{\mathbf{R}}_{m,k}$ into the *a posteriori* probability

$$p(\mathbf{X}_k, \widehat{\mathbf{s}}_{m,k}, \widehat{\mathbf{R}}_{m,k}, H_{m,k}) = p(\mathbf{X}_k|\widehat{\mathbf{s}}_{m,k}, \widehat{\mathbf{R}}_{m,k}, H_{m,k})$$
$$\Pr(\widehat{\mathbf{s}}_{m,k}|\widehat{\mathbf{R}}_{m,k}, H_{m,k})\, p(\widehat{\mathbf{R}}_{m,k}|H_{m,k})\, \Pr(H_{m,k}) \quad (11)$$

and then maximizing $p(\mathbf{X}_k, \widehat{\mathbf{s}}_{m,k}, \widehat{\mathbf{R}}_{m,k}, H_{m,k})$ over all M possible hypotheses to determine the optimal track associations. Note that the 1st three terms on the right in (11) incorporate the joint dependence of the slant tracks on the raymodes and on the targets' ground locations, the dependence of the raymodes over ground range, and the *a priori* ground coordinate information under each hypothesis. For a single revisit, the *a priori* hypothesis probabilities are $P(H_{m,k}) = 1/M$ for $m = 1$ to M.

To provide consistent mode linking decisions over multiple revisits, the method is combined with a simple model of the temporal correlation between mode linking hypotheses. To reduce computation, the slant track observations, raymode estimates and ground track estimates are modeled as independent from revisit to revisit. This assumption avoids additional target dynamics constraints beyond those resulting from tracking done in slant coordinates. The mode linking hypotheses over multiple revisits are modeled as a first-order Markov chain to limit the GLRT solutions to track associations that are consistent from revisit to revisit. The MAP mode linking hypothesis sequence estimate for k revisits is then given by

$$[\widehat{H}_1, ..., \widehat{H}_k] = \arg\max_{H_1,...,H_k} \sum_{l=1}^{k} (\log p(X_k, \widehat{\mathbf{s}}_{m,l}, \widehat{\mathbf{R}}_{m,l}|H_{m,l})$$
$$+ \log \Pr(H_{m,l}|H_{m,l-1})) \quad (12)$$

where $p(X_k, \widehat{\mathbf{s}}_{m,l}, \widehat{\mathbf{R}}_{m,l}|H_{m,l})$ is the product of the 1st 3 terms on the right in (11) and $\Pr(H_{m,l}|H_{m,l-1})$ is non-zero only for hypothesis pairs with consistent slant-track-to-target assignments. This maximization can be performed via a fast recursive method that grows linearly with the number of revisits. This is particularly important in the case of mode linking where the number of hypothesis states grows exponentially with the number of slant tracks at each revisit. The MAP mode linking approach optimally weights the time history of mode linking decisions with the current slant track data. In contrast, conventional mode linking approaches often make a hard decision after a limited observation time.

C. Real Data Results

To illustrate the implementation of the MAP mode linking method, a flow diagram is shown in Figure 7. The inputs consist of just the ionospheric model statistics and the slant track observations. The estimates include the MAP mode linking

108
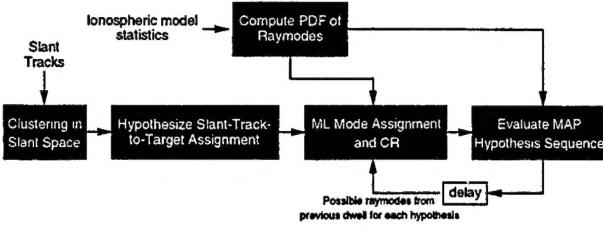
Fig. 7. MAP mode linking block diagram

hypothesis sequence, the targets' ground tracks and the ray-mode assignments for all the slant tracks. Given the ionospheric statistics that consist of the plasma frequency profile parameters' mean and covariance, the PDF of the raymodes is computed via Monte Carlo raytracing. In order to reduce computation due to the inclusion of extremely unlikely combinations, the slant track observations over multiple revisits are first clustered according to their relative Doppler frequencies, slant azimuths and slant azimuth rates. All possible slant-track-to-target assignments are then enumerated using a hypothesis tree like the one shown in figure 6. Given the slant track data and the statistical propagation model results, each branch of the hypothesis tree is then evaluated at the kth revisit as follows. The MLCR method is used to estimate the raymode assignments and the location of each individual target under each mode linking hypothesis. The MAP mode linking hypothesis decision is then determined based on the data, the MLCR estimates, and the history of previous mode linking decisions. To constrain each slant track's raymode assignments from revisit to revisit, a list of all the possible raymodes, for the ground location estimate, is retained for each slant track under each hypothesis until the next revisit. The lists are then used at the next revisit to limit the raymode types to only those that are consistent with the slant tracks' previous ground track location estimates under the same mode linking assignment.

The performance of the MAP mode linking approach was compared to that of a conventional mode linker of an existing OTH radar (OTHR) on roughly 90 minutes (180 revisits) of real OTH radar slant track data from dwell illumination region (DIR) 200 between 1946 and 2111 UT on Sep. 22, 1998. Figure 8 shows the slant ranges versus time of all the slant tracks generated from DIR 200. Each slant track is indicated by a different
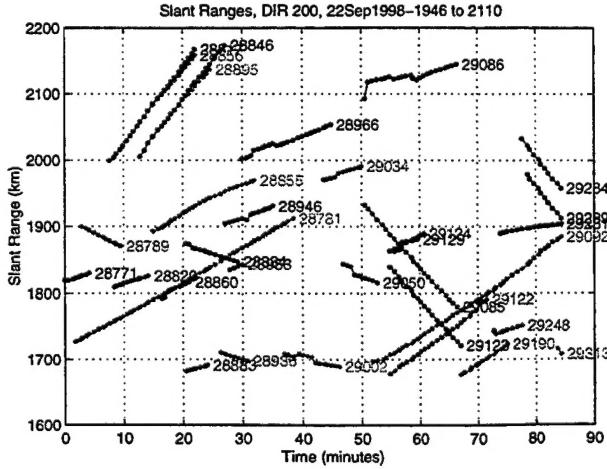

Fig. 8. OTHR slant tracks: DIR 200, Sep. 22, 1998, 1946:2111 UT

number ID. Several potential multipath cases are clearly evident

from the nearly parallel slant tracks with similar initiation times and durations. In particular in figure 8, note the potential multipath scenario involving slant tracks 28817, 28846, 28856 and 28895 in the upper left and also the scenario with slant tracks 29284 and 29289 midway up on the right.

Figure 9 compares the ground tracks of the conventional OTHR mode linking approach and the ground tracks of the MAP mode linking approach to FAA ground truth data. The conventional mode linking approach is to evaluate of all possible raymode combinations and then link slant tracks together whenever their corresponding ground tracks have sufficiently minimum variance. Since this may result in unrealisitic ground tracks with jagged jumps in the tracks, a jump constraint is employed between revisits to provide smoother ground tracks. The line symbols indicate the target positions at each different revisit. The conventional OTHR mode linking ground tracks are indicated by diamond line symbols, the MAP mode linking ground tracks are indicated by the circles and the FAA ground tracks are indicated by the pentagram line symbols. Note that


(a)

(b)
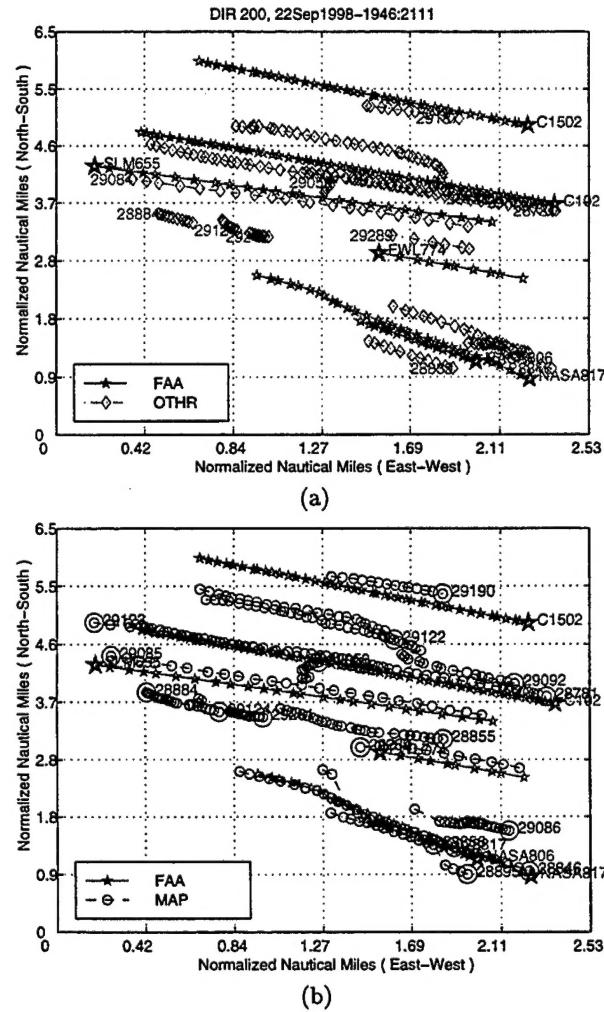
Fig. 9. Comparisons with FAA ground truth tracks for slant tracks from DIR 200 at 22Sep1998-1946:2111 UT using a) conventional mode linking (OTHR) and b) MAP mode linking

in terms of their general character, the ground tracks of the MAP mode linking approach are very similar to the conventional mode linker's ground tracks. In both cases, the ground tracks are quite smooth from revisit to revisit and the number of estimated ground tracks is close to the number of FAA ground

tracks.

Figure 10 shows a more detailed ground track geographic display of the multi-target multi-track scenario involving slant tracks 28817, 28846, 28856, and 28895. Both the conventional
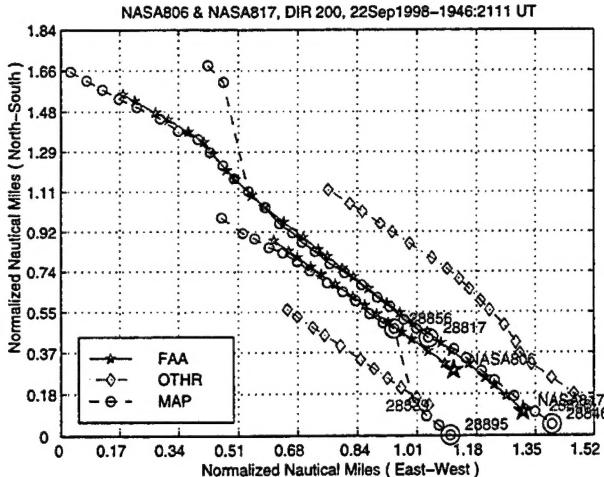

Fig. 10. Flights NASA806 and NASA817 on 22Sep1998

mode linker and the MAP mode linker show two ground tracks, but the MAP ground tracks show a significant improvement in track accuracy. The MAP mode linker assigned slant tracks 28817 and 28846 to the same target (NASA817) on raymodes F2L-F2L and E-F2L respectively with a median absolute miss distance of 0.09688 in normalized nautical miles (n-nmi). The conventional mode linker also assigned 28817 and 28846 to the same target with both on raymode F2L-F2L and the median miss distance was 0.23438 n-nmi. The MAP mode linking approach assigned tracks 28856 and 28895 to the other target (NASA806) on raymodes F2L-F2L and E-F2L with median miss distances of 0.09375 and 0.15469 n-nmi respectively. The conventional mode linker formed a ground track for only 28856 on raymode F2L-F2L with a median miss distance of 0.28750 n-nmi while slant track 28895 was not put to ground by the conventional OTHR mode linker. In this scenario, that MAP mode linking approach provides as much as a 3:1 accuracy improvement over conventional mode linking.

Figure 11 shows a close-up view of the single target multi-track scenario involving slant tracks 29284 and 29289. Note
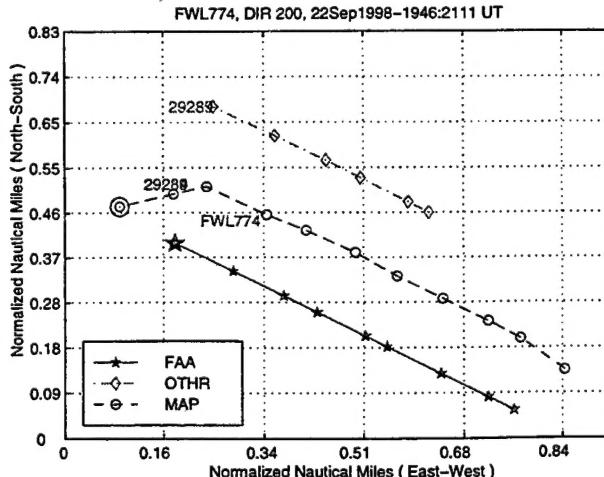

Fig. 11. Flight FWL774, DIR 200, 22Sep1998-1946:2111 UT

that the MAP mode linker's ground track is much closer than the conventional mode linker's ground track to the true ground track of flight FWL774. The MAP mode linker assigned the slant tracks 29284 and 29289 to the same target and to raymodes F2L-F2L and F1L-F2L giving a median miss distance of 0.15625 n-nmi. The conventional mode linker assigned 29284 and 29289 to two different targets and both to raymode F2L-F2L. Only one OTHR ground track is shown in figure 11 since 29289 was linked to another slant track from an overlapping DIR that was not used here. The median miss distance of OTHR ground track 29284 is 0.28907 n-nmi. In this case, the MAP mode linker provides nearly a 2:1 accuracy improvement over the conventional OTHR mode linker.

IV. CONCLUSIONS

In this paper, matched-field altitude estimation and maximum a posteriori mode linking for over-the-horizon radar have been presented. These techniques represent just two examples of how complex multipath propagation conditions can be modeled and exploited to provide, respectively, a new capability and enhanced performance to an existing radar. Clearly, such approaches involve a tight coupling between the physics of wave propagation and signal processing. They also rely on the availability of sufficiently accurate estimates of the environmental parameters. Numerous results obtained with real data in very different settings, however, suggest that "sufficiently accurate" should by no means be interpreted as "perfect knowledge" of the environment.

V. ACKNOWLEDGEMENT

REFERENCES

[1] J. M. Headrick and M. I. Skolnik, "Over-the-horizon radar in the HF band," *Proceedings of the IEEE*, vol. 62, pp. 664–673, June 1974.

[2] L. F. McNamara, *The Ionosphere: Communications, Surveillance, and Direction Finding*. Orbit, a foundation series, Malabar, FL: Krieger Publishing Co., 1991.

[3] J. L. Krolik and R. H. Anderson, "Maximum likelihood coordinate registration for over-the-horizon radar," *IEEE Transactions on Signal Processing*, vol. 45, pp. 945–959, April 1997.

[4] A. B. Baggeroer, W. A. Kiperman, and H. Schmidt, "Matched-field processing: Source localization in correlated noise as an optimum parameter estimation problem," *Journal of the Acoustical Society of America*, vol. 83, pp. 571–587, February 1988.

[5] J. K. Jao, "A matched array beamforming technique for low angle radar tracking in multipath," in *IEEE National Radar Conference*, pp. 171–176, 1994.

[6] D. F. Gingras, P. Gerstoft, and N. L. Gerr, "Electromagnetic matched-field processing: basic concepts and tropospheric simulations," *IEEE Transactions on Antennas and Propagation*, vol. 45, pp. 1536–1545, October 1997.

[7] M. Papazoglou and J. L. Krolik, "Electromagnetic matched-field processing for over-the-horizon radar in an uncertain environment," *Submitted to IEEE Transactions on Antennas and Propagation*, May 1997.

[8] T. A. Croft and H. Hoogasian, "Exact ray calculations in a quasi-parabolic ionosphere with no magnetic field," *Radio Science*, vol. 3, pp. 69–74, January 1968.

[9] E. J. Ferraro, *ROTHR Mode Linking*. Raytheon Electronic Systems, August 1995. Memo Number EJF:95:13.

[10] R. H. Anderson and J. L. Krolik, "Over-the-horizon radar target localization using a hidden markov model estimated from ionosonde data," *Radio Science*, vol. 33, pp. 1199–1213, July-August 1998.

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 9 July 1999	3. REPORT TYPE AND DATES COVERED Project Report ASAP-7, Volume 1

4. TITLE AND SUBTITLE

Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop, 10–11 March 1999

5. FUNDING NUMBERS

C — F19628-95-C-0002
PR — 682

6. AUTHOR(S)

Cindy J. Pelley, Editor

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Lincoln Laboratory, MIT
244 Wood Street
Lexington, MA 02173-9108

8. PERFORMING ORGANIZATION REPORT NUMBER

ASAP-7, Volume 1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

DARPA
3701 N. Fairfax Drive
Arlington, VA 22203-1714

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ESC-TR-98-054

11. SUPPLEMENTARY NOTES

None

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This year marks the seventh annual ASAP workshop, which is sponsored jointly by the DARPA Sensor Technology and Tactical Technology Offices, the Navy AEGIS and E2C Program Offices, and the NAVSEA Advanced Systems and Technology Office. This joint sponsorship reflects the far-reaching implications of adaptive sensor array processing for a wide range of military applications.

As this workshop has evolved over its seven year history, a common theme has been to highlight the similarity of adaptive sensor processing techniques for various disciplines. The ASAP forum has allowed researchers from academia, government, and industry to discuss common problems, developing ideas, and solutions related to how these diverse technologies can be applied to national defense interests. This year will continue to stress the strong coupling between the radar, sonar, and communications communities. A key theme will be the system implications of adaptive signal processing techniques. It is hoped that the cross-fertilization of ideas at ASAP will provide new areas of exploration and accelerate technological advances benefitting national defense interests.

14. SUBJECT TERMS

15. NUMBER OF PAGES
114

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Same as Report	19. SECURITY CLASSIFICATION OF ABSTRACT Same as Report	20. LIMITATION OF ABSTRACT Same as Report